Box Seq. A

### JM PTO-1082 46 U.s. p

HOWREY & SIMON
Box No. 34
1299 Pennsylvania Avenue, N.W.
Washington, D.C. 20004-2402
(650) 463-8100

Attorney Docket No. 5371.31.US02

**Box Patent Application**ASSISTANT COMMISSIONER FOR PATENTS
Washington, D.C. 20231

Sir:

Transmitted herewith for filing is the patent application of Patricia D. Murphy, Marga B. White, Mark B. Rabin; Sheri J. Olson; Matthew Yoshikawa; Geoffrey M. Jackson; Tara Eskandari; Brenda Schryer; and Michael Park for NOVEL CODING SEQUENCE HAPLOTYPES OF THE HUMAN BRCA2 GENE.

Also, enclosed are:

- 1. Cover Sheet of Application;
- 2. 13 sheets of drawings;
- 3. Executed Declaration (2 sets);
- 4. Executed Small Entity Statement;
- 5. Assignment Cover Sheet;
- 6. Executed Assignment (2 sets);
- 7. Sequence Listing on Disk; and
- 5. Two (2) return postcards.

The filing fee has been calculated as shown below:

•	(Col. 1)		(Col. 2)
FOR	NO. FILED	NC	). EXTRA
BASIC FEE			
TOTAL CLAIMS	60-20 =	*	40
INDEP. CLAIMS	29 -3 =	*	26
■ MULTIPLE DEP	ENDENT CLAIN	1 PR	ESENTED

<sup>\*</sup>If the difference in Col. 1 is less than zero, enter "0" in Col. 2

SMALL	ENTITY _	
RATE	FEE	
	\$ 395.00	
40 X 11 =	440.00	
26 x 41 =	1,066.00	
+ 135 =	135.00	
TOTAL	\$2,036.00	

	011121	THAN A LENTITY
OR	RATE	FEE
OR		\$ 790.00
OR	X 22 =	
OR	X 82=	
OR	- 270 =	
OR	TOTAL	
]		

A check in the amount of \$2,076.00 (\$2,036.00 filing fee + \$40.00 recording of Assignment) is enclosed.

The U.S. Patent and Trademark Office is hereby authorized to charge any fee deficiency, or credit any overpayment, to our Deposit Account No. <u>08-3038</u>. A duplicate copy of this sheet is attached.

Date May 22, 1998

Albert P. Halluin (Reg. No.25,227)

Ξ

Name (Print/Type)

Signature

### UTILITY PATENT APPLICATION

Attorney	Docket No.	5371.31.US02	
First Nam	ned Inventor or A	Application Identifier	Patricia D. Murphy
Title	Novel Coding	Sequence Haplotypes of	of the Human BRCA2

TRANSMITTAL fer nonprovisional applications under 37 CFR 1 53(h)) Express Mail Label No. EM555262526US APPLICATION ELEMENTS Assistant Commissioner for Patents See MPEP chapter 600 concerning utility patent application contents ADDRESS TO: Box Patent Application Washington, DC 20231 \*Fee Transmittal Form (Form PTO-1082) 6. Microfiche Computer Program (Appendix) (Submit an original and a duplicate for fee processing) Specification 2. [Total Pages | 204 Nucleotide and/or Amino Acid Sequence Submission (preferred arrangement set forth below) (if applicable, all necessary) - Descriptive title of the Invention Computer Readable Copy - Cross References to Related Applications - Statement Regarding Fed sponsored R&D Paper Copy (identical to computer copy) b. - Reference to Microfiche Appendix - Background of the Invention Statement verifying identity of above c. - Brief Summary of the Invention copies - Brief Description of the Drawings (if filed) ACCOMPANYING APPLICATION PARTS Assignment Papers (cover sheet & document(s)) 8. - Detailed Description 9. - Claims 37 CFR 3.73(b) Statement Power of - Abstract of the Disclosure (when there is an assignee) Attorney [Total Sheets 13 Drawing(s) (35 USC 113) 10. English Translation Document (if applicable) [Total Pages | 6 Oath or Declaration Information Disclosure Copics of IDS 11. Statement (IDS)/PTO-1449 Citations -1 Newly executed (original or copy) a. 12. Preliminary Amendment Copy from a prior application (37 CFR 1.63(d)) Return Receipt Postcard (MPEP 503) (Two) b. 13. (for continuation/divisional with Box 17 completed) (should be specifically itemized) [Note Box 5 below] **DELETION OF INVENTOR(S)** 14. \*Small Entity Statement filed in Signed statement attached deleting inventor(s) named Statement(s) prior application, in the prior application, see 37 CFR 1 63(d)(2) and Status still proper 1.33(b) and desired 15. Certified Copy of Priority Document(s) (if foreign priority is claimed) Incorporation By Reference (useable if Box 4b is checked) 16. Other: The entire disclosure of the prior application, from which a copy of the oath or declaration is supplied under Box 4b, is considered \*NOTE FOR ITEMS 1 & 14 IN ORDER TO BE ENTITLED TO PAY SMALL ENTITY FEFS, A SWALL ENTITY STATEMENT IS REQUIRED (37 C F R  $\S$  1 27), EXCEPT IF ONE FILED IN A PRIOR APPLICATION IS RELIED UPON (37 C F R  $\S$  1 28) as being part of the disclosure of the accompanying application and is hereby incorporated by reference therein. 17. If a CONTINUING APPLICATION, check appropriate box and supply the requisite information: Divisional Continuation-in-part (CIP) of prior application No Prior Application Information: Examiner: Group/Art Unit: 18. CORRESPONDENCE ADDRESS Customer Number or Bar Code Label or \( \sum \) Correspondence address below (Insert Customer No or Attach bar code label here) Albert P. Halluın HOWREY & SIMON NAME Box No. 34 ADDRESS 1299 Pennsylvania Avenue, N.W  $\overline{CITY}$ Washington STATE DC ZIP CODE 20004-2402 COUNTRY US **TELEPHONE** 202-783-0800 FAX 202-383-7195

Registration No (Attorney/Agent)

25

May 22,

1998

### APPLICATION IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

### **FOR**

### NOVEL CODING SEQUENCE HAPLOTYPES OF THE HUMAN BRCA2 GENE

### **INVENTORS**:

Patricia D. Murphy

U.S. citizen

Slingerlands, NY

Marga B. White

U.S. citizen

Frederick, MD

Mark B. Rabin

U.S. citizen

Rockville, MD

Sheri J. Olson

U.S. citizen

Falls Church, VA

Matthew Yoshikawa

U.S. citizen

Germantown, MD

Geoffrey M. Jackson

U.S. citizen

Beltsville, MD

Tara Eskandari

U.S. citizen

Rockville, MD

Brenda Schryer

U.S. citizen

Bel Air, MD

Michael Park

U.S. citizen

Rocksville, MD

HOWREY & SIMON 1299 Pennsylvania Ave., N.W. Box 34 Washington, D.C. 20004 (650) 463-8100

Attorney Docket No. 5371.31.US02

### NOVEL CODING SEQUENCE HAPLOTYPES OF THE HUMAN BRCA2 GENE

This is an U.S. utility patent application based on U.S. Provisional Application Serial Nos. 60/055,784 filed on August 15, 1997, 60/064,926 filed on November 7, 1997, and 60/065,367 filed on November 12, 1997.

### FIELD OF THE INVENTION

This invention relates to a gene which has been associated with breast cancer where the gene is found to be mutated. More specifically, this invention relates to five unique coding sequences of BRCA2 gene BRCA2<sup>(omi1)</sup>, BRCA2<sup>(omi2)</sup>, BRCA2<sup>(omi3)</sup>, BRCA2<sup>(omi4)</sup>, and BRCA2<sup>(omi5)</sup> identified in human subjects which define five novel haplotypes.

### **BACKGROUND OF THE INVENTION**

It has been estimated that about 5-10% of breast cancer is inherited (Rowell, S., et al., American Journal of Human Genetics 55:861-865 (1994)). The first gene associated with both breast and ovarian cancer was cloned in 1994 from chromosome 17 by Miki, Y., et al., Science 266:66-71 (1994). A second high-risk breast cancer conferring gene was located on chromosome 13 in 1994 (Wooster, R., et al., Science 265:2088-2090) and subsequently cloned in 1995 (Wooster, R., et al., Nature 378:789-792). Mutations in this "tumor suppressor" gene are thought to account for roughly 35% of inherited breast cancer and 80-90% of families with male breast cancer.

Locating one or more mutations in the BRCA2 region of chromosome 13 provides a promising approach to reducing the high incidence and mortality associated with breast cancer through the early detection of women and men at high risk. These individuals, once identified, can be targeted for more aggressive prevention programs. Screening is carried out by a variety of methods which include karyotyping, probe binding and DNA sequencing.

In DNA sequencing technology, genomic DNA is extracted from whole blood and the coding regions of the BRCA2 gene are amplified. Each of the coding regions may be sequenced completely and the results are compared to the normal DNA sequence of the gene. Alternatively, the coding sequence of the sample gene may be compared to a panel of known mutations or other screening procedure

before completely sequencing the gene and comparing it to a normal sequence of the gene.

The BRCA2 gene is divided into 27 separate exons. Exon 1 is noncoding, in that it is not part of the final functional BRCA2 protein product. The BRCA2 coding region spans roughly 10433 base pairs (bp) over 70 kb. Each exon consists of 100-600 bp, except for exons 10, 11 and 27. The full length mRNA is 11-12 kb. To sequence the coding region of the BRCA2 gene, each exon is amplified separately and the resulting PCR products are sequenced in the forward and reverse directions. Because exons 10, 11, and 27 are so large, we have divided them into three, twenty-one, and two overlapping PCR fragments (respectively) of approximately 250-625 bp each (segments "A" through "C" of exon 10, "A" through "U" of exon 11, and "A" through "B" of exon 27).

Many mutations and normal polymorphisms have already been reported in the BRCA2 gene. A world wide web site has been built to facilitate the detection and characterization of alterations in breast cancer susceptibility genes. Such mutations in BRCA2 can be accessed through the Breast Cancer Information Core (BIC) at http://www.nhgri.nih.gov/Intramural\_research/Lab\_transfer/Bic. This data site became publicly available on November 1, 1995. Friend, S. et al. Nature Genetics 11:238, (1995). The information on BRCA2 was added in February, 1996.

The genetics of Breast Cancer Syndrome is autosomal dominant with reduced penetrance. In simple terms, this means that the syndrome runs through families: (1) both sexes can be carriers (mostly women get the disease but men can both pass it on and occasionally get breast cancer); (2) most generations will likely have breast cancer; (3) occasionally women carriers either die young before they have the time to manifest disease (and yet have offspring who get it) or they never develop breast or ovarian cancer and die of old age (the latter people are said to have "reduced penetrance" because they never develop cancer). Pedigree analysis and genetic counseling is absolutely essential to the proper workup of a family prior to any lab work.

Until now, the only sources of genomic sequence information for BRCA2 were GenBank (Accession Number U43746), or through the Breast Information Core (BIC) database on the Internet which requires membership in the BIC consortium. However, based upon the disclosure of this patent application, in neither GenBank

nor BIC were the sequences identified and listed entirely accurate. There is a need in the art to correct these mistakes which otherwise may lead to misinterpretation of the sequence data from the patient as abnormal when it was not, or vice versa.

In addition, there is a need in the art to have available a functional allele profile which represents the most likely BRCA2 sequences to be found in the majority of the normal population. This functional allele profile is based upon frequent polymorphisms and the correct backbone sequence. The knowledge of several common normal haplotypes will make it possible for true mutations to be easily identified or differentiated from polymorphisms. Identification of mutations of the BRCA2 gene and protein would allow more widespread diagnostic screening for hereditary breast cancer than is currently possible.

The use of these common normal haplotypes, in addition to the previously published BRCA2 sequence, will reduce the likelihood of misinterpreting a "sequence variation" found in the normal population with a pathologic "mutation" (i.e. causes disease in the individual or puts the individual at a high risk of developing the disease). With large interest in breast cancer predisposition testing, misinterpretation is particularly worrisome. People who already have breast cancer are asking the clinical question: "is my disease caused by a heritable genetic mutation?" The relatives of the those with breast cancer are asking the question: "Am I also a carrier of the mutation my relative has? Thus, is my risk increased, and should I undergo a more aggressive surveillance program?"

### **SUMMARY OF THE INVENTION**

The present invention is based on the discovery of the correct genomic BRCA2 sequence and five novel sequence haplotypes found in normal human subjects of the BRCA2 gene.

It is an object of this invention to provide the correct intronic/exonic sequence of the BRCA2 gene.

It is another object of this invention to provide five unique haplotype sequences of the BRCA2 gene in normal individuals which do not correspond to increased cancer susceptibility.

It is another object of this invention to sequence a BRCA2 gene or a portion thereof and compare it to the five haplotype sequences to determine whether a

sequence variation noted represents a polymorphism or a potentially harmful mutation.

It is another object of this invention to provide a list of the pairs which occur at each of ten polymorphic points in the BRCA2 gene.

It is another object of this invention to provide the rates of occurrence for the polymorphisms at codons 289, 372, 455, 743, 894, 991, 1132, 1269, 2414, and 2951 in the BRCA2 gene.

It is another object of this invention to provide a method wherein all exons of BRCA2 gene or parts thereof, are amplified with one or more oligonucleotide primers.

It is another object of this invention to provide a method of identifying a individual who carries no mutation(s) of the BRCA2 gene and is therefore at no increased risk or susceptibility to breast or ovarian cancer based on a finding that the individual does not carry an abnormal BRCA2 genes.

It is another object of this invention to provide a method of identifying a mutation in BRCA2 gene leading to predisposition or higher susceptibility to breast or ovarian cancer.

It is another object of this invention to provide five novel BRCA2 protein sequences derived from five BRCA2 haplotype sequences.

It is another object of the invention to encompass prokaryotic or eukaryotic host cells comprising an expression vector having a DNA sequence that encodes for all or a fragment of the five novel BRCA2 protein sequences, a BRCA2 polypeptide thereof, or a functional equivalent thereof.

It is another object of the invention to encompass an anti-BRCA2 protein antibody using all of fragments of the five novel BRCA2 protein sequences, a BRCA2 polypeptide thereof or a functional equivalent thereof as an immunogen.

There is a need in the art for cDNA sequences of the BRCA2 gene and for the protein sequences of BRCA2 gene from normal individuals who are not at risk for increased susceptibility for cancer. In order to determine whether a sample from a patient suspected of containing a BRCA2 mutation actually has the mutation, the patient's BRCA2 DNA and/or amino acid sequence need to be compared to all known normal BRCA2 sequences. Failure to compare the sequence obtained to all

naturally occurring normal sequences may result in reporting a sample as containing a potentially harmful mutation when it is a polymorphism without clinical significance.

A person skilled in the art of genetic susceptibility testing will find the present invention useful for:

- identifying individuals having a normal BRCA2 gene with no coding sequence mutations, who therefore cannot be said to have an increased genetic susceptibility to breast or ovarian cancer from their BRCA2 genes;
- avoiding misinterpretation of normal polymorphisms found in the BRCA2 gene;
- determining the presence of a previously unknown mutation in the BRCA2 gene;
- d) identifying a mutation in exon 11 of BRCA2 which indicates a predisposition or higher susceptibility to ovarian cancer than breast cancer (i.e., resides in the putative "ovarian cancer cluster" region);
- e) probing a human sample of the BRCA2 gene by allele to determine the presence or absence of either polymorphic alleles or mutations;
- f) performing gene therapy with the correct BRCA2 gene sequence.
- g) performing protein replacement therapy with the correct BRCA 2 protein sequence or a functional equivalent thereof.

### **BRIEF DESCRIPTION OF THE FIGURES**

FIGURE 1 shows the GenBank genomic sequence of BRCA2 (Accession Number U43746). The lower case letters denote intronic sequences and the upper case letters denote exonic sequences. Incorrect exonic sequences at exons 5 and 16 are shown with boldface type.

FIGURE 2 shows the corrected genomic sequence of BRCA2. The lower case letters denote intronic sequences and the upper case letters denote exonic sequences. Corrected intronic and exonic sequences at exons 5, 11 and 15 are shown with boldface type.

FIGURE 3 shows the alternative alleles at polymorphic sites along a chromosome which can be represented as a unit or "haplotype" within a gene such as BRCA2.

The haplotype that is in GenBank (GB) is shown with light shading. Five additional haplotypes are shown in FIGURE 3 (encompassing the alternative alleles found at nucleotide sites 1093, 1342, 1593, 2457, 2908, 3199, 3624, 4035, 7470 and 9079). BRCA2 (omi-1), BRCA2 (omi-2), BRCA2 (omi-3), BRCA2 (omi-4), and BRCA2 (omi-5) are represented with mixed dark and light shading (numbers 2, 4, 6, 8 and 10 from left to right). In total, 5 of 10 haplotypes along the BRCA2 gene are unique.

### **DETAILED DESCRIPTION OF THE INVENTION**

### **DEFINITIONS**

The following definitions are provided for the purpose of understanding this invention.

"Breast and Ovarian cancer" is understood by those skilled in the art to include breast, ovarian and pancreatic cancer in women and also breast, prostate and pancreatic cancer in men. BRCA2 is associated with genetic susceptibility to breast, ovarian and pancreatic cancer. Therefore, claims in this document which recite breast and/or ovarian cancer refer to breast, ovarian, prostate, and pancreatic cancers in men and women.

"Coding sequence" refers to those portions of a gene which, taken together, code for a peptide (protein), or which nucleic acid itself has function.

"Protein" or "peptide" refers to a sequence of amino acids which has function.

"BRCA2<sup>(omi)</sup>" refers to the genomic BRCA2 sequence disclosed in Genbank (Accession Number U43746) wherein,

- (1) a 10 bp stretch (5'-TTTATTTTAG-3') is intronic at 3' end of intron 4, rather than at the 5' end of exon 5; and
- (2) a 16 bp stretch (5'-GTGTTCTCATAAACAG-3') is exonic at the 3' end of exon 15, rather than at the 5' end of exon.

"BRCA2<sup>(omi 1-5)</sup>" refers to five unique DNA sequences of the BRCA2 gene and their introns (particularly the slice sites adjacent to the exons). These sequences were found by end to end sequencing of the BRCA2 gene from 5 individuals randomly drawn from the population and who were documented to have no family history of breast or ovarian cancer. The sequenced exons were found not to contain any truncating mutations. In all cases the change of a nucleic acid at a

polymorphic site lead to a codon change and a change of amino acid from the previously published standard in GenBank (see TABLE III). In some cases the frequency of occurrence of a nucleic acid change was found to differ from the published frequency or was newly determined. These sequence variations are believed to be alleles whose haplotypes do not indicate an increased risk for cancer.

"Normal DNA sequence" also called "normal gene sequence" refers to a nucleic acid sequence, the nucleic acid of which are known to occur at their respective positions with high frequency in a population of individuals who carry the gene which codes for a normally functioning protein, or which itself has normal function.

"Normal Protein Sequence" refers to the protein sequence, the amino acids of which are known to occur with high frequency in a population of individuals who carry the gene which codes for a normally functioning protein.

"Normal Sequence" refers to the nucleic acid or protein sequence, the nucleic or amino acids of which are known to occur with high frequency in a population of individuals who carry the gene which codes for a normally functioning protein, or which nucleic acid itself has a normal function.

"Haplotype" refers to a series of specific alleles within a gene along a chromosome.

"Functional allele profile" refers a list of those alleles in the normal population which have the funll function.

"Mutation" refers to a base change or a gain or loss of base pair(s) in a DNA sequence, which results in a DNA sequence coding for a non-functional protein or a protein with substantially reduced or altered function.

"Polymorphism" refers to a base change in a DNA sequence which is not associated with known pathology.

"Primer" refers to a sequence comprising about 15 or more nucleotides having a sequence complementary to the BRCA2 gene. Other primers which can be used for primer hybridization will be known or readily ascertainable to those skilled in the art.

"Substantially complementary to" refers to primer sequences which hybridize to the sequences provided under stringent conditions and/or sequences having sufficient homology with BRCA2 sequences, such that the allele specific oligonucleotide primers hybridize to the BRCA2 sequences to which they are complimentary.

"Isolated nucleic acids" refers to nucleic acids substantially free of other nucleic acids, proteins, lipids, carbohydrates or other materials with which they may be associated. Such association is typically either in cellular material or in a synthesis medium.

"Biological sample" or "body sample" refers to a sample containing DNA oatained from a biological source. The sample may be from a living, dead or even archeological source from a variety of tissues and cells. Examples include body fluid (e.g. blood (leukocytes), urine (epithelial cells), saliva, breast milk, menstrual flow, cervical and vaginal secretions, etc.), skin, hair roots/follicle, mucus membrane (e.g. buccal or tongue cell scrapings), cervicovaginal cells (from PAP smear, etc.), lymphatic tissue, internal tissue (normal or tumor).

"Vector" refers to any polynucleotide which is capable of self replication or inducing integration into a self-replicating polynucleotide. Examples include polynucleotides containing an origin or replication or an integration site. Vectors may be intergrated into the host cell's chromosome or form an autonomously replicating unit.

"A tumor growth inhibitor" refers to a molecule such as, all or a fragment of BRCA2 protein, a BRCA2 polypeptide, or a functional equivalent thereof that is effective for preventing the formation of, reducing, or eliminating a transformed or malignant phenotype of breast or ovarian cancer cells.

"A BRCA2 polypeptide" refers to a BRCA2 polypeptide either directly derived from the BRCA2 protein, or homologous to the BRCA2 protein, or a fusion protein consisting of all or fragments of the BRCA2 protein and polypeptides.

"A functional equivalent" refers to a molecule including an unnatural BRCA2 polypeptide, a drug or a natural product which retains substantial biological activity as the native BRCA2 protein. The activity and function of BRCA2 protein may include transactivation, granin, DNA repair, among others.

"A target polynucleotide" refers to the nucleic acid sequence of interest, for example, the BRCA2 encoding polynucleotide. Other primers which can be used for primer hybridization will be known or readily ascertainable to those of skill in the art.

The invention in several of its embodiments includes: an isolated DNA sequence of the BRCA2 coding sequence as set forth in SEQ ID NO:4, 6, 8, 10, and 12, a protein sequence of the BRCA2 protein as set forth in SEQ ID NO:5, 7, 9, 11, 13, a method of identifying individuals having a normal BRCA2 gene with no increased risk for breast and ovarian cancer, a method of detecting an increased genetic susceptibility to breast and ovarian cancer in an individual resulting from the presence of a mutation in the BRCA2 coding sequence, a method of performing gene therapy to prevent or treat a tumor, a method of protein replacement therapy to prevent or treat a tumor, a diagnostic reagent comprising all or fragments of the disclosed BRCA2 cDNA and protein sequences.

### SEQUENCING

Any nucleic acid specimen, in purified or non-purified form, can be utilized as the starting nucleic acid, providing it contains, or is suspected of containing, the specific nucleic acid sequence containing a polymorphic or a mutant allele. Thus, the process may amplify, for example, DNA or RNA, including mRNA and cDNA, wherein DNA or RNA may be single stranded or double stranded. In the event that RNA is to be used as a template, enzymes and/or conditions optimal for reverse transcribing the template to DNA would be utilized. In addition, a DNA-RNA hybrid which contains one strand of each may be utilized. A mixture of nucleic acids may also be employed, or the nucleic acids produced in a previous method such as an ... amplification reaction using the same or different primers may be so utilized. The specific nucleic acid sequence to be amplified, i.e., the polymorphic and/or the mutant allele, may be a fraction of a larger molecule or can be present initially as a discrete molecule, so that the specific sequence constitutes the entire nucleic acid. A variety of amplification techniques may be used such as ligating the DNA sample or fragments thereof to a vector capable of replication or incorporation into a replicating system thereby increasing the number of copies of DNA suspected of containing at least a portion of the BRCA2 gene. Amplification techniques include so called "shot gun cloning". It is not necessary that the sequence to be amplified be present initially in a pure form; it may be a minor fraction of a complex mixture, such as contained in whole human DNA.

It should be noted that one need not sequence the entire coding region or even an entire DNA fragment in order to determine whether or not a mutation is present. For example, when a mutation is known in one family member, it is sufficient to determine the sequence at only the mutation site by sequencing or by other mutation detection systems such as ASO when testing other family members.

DNA utilized herein may be extracted from a body sample, such as blood, tissue material and other biological sample by a variety of techniques such as that described by Maniatis, et al. in Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, NY, p 280-281, 1982). If the extracted sample is impure, it may be treated before amplification with an amount of a reagent effective to open the cells, and to expose and/or separate the strand(s) of the nucleic acid(s). This lysing and nucleic acid denaturing step to expose and separate the strands will allow amplification to occur much more readily.

For amplification by cloning, the isolated DNA may be cleaved into fragments by a restriction endonuclease or by shearing by passing the DNA containing mixture through a 25 gauge needle from a syringe to prepare 1-1.5 kb fragments. The fragments are then ligated to a cleaved vector (virus, plasmid, transposon, cosmid etc.) and then the recombinant vector so formed is then replicated in a manner typical for that vector.

For a PCR amplification, the deoxyribonucleotide triphosphates dATP, dCTP, dGTP, and dTTP are added to the synthesis mixture, either separately or together with the primers, in adequate amounts and the resulting solution is heated to about 90°-100°C from about 1 to 10 minutes, preferably from 1 to 4 minutes. After this heating period, the solution is allowed to cool, which is preferable for the primer hybridization. To the cooled mixture is added an appropriate agent for effecting the primer extension reaction (called herein "agent for polymerization"), and the reaction is allowed to occur under conditions known in the art. The agent for polymerization may also be added together with the other reagents if it is heat stable. This synthesis (or amplification) reaction may occur at room temperature up to a temperature above which the agent for polymerization no longer functions. Thus, for example, if DNA polymerase is used as the agent, the temperature is generally no greater than about 40°C. Most conveniently the reaction occurs at

room temperature. When using thermostable DNA polymerase such as Taq, higher temperature may be used.

The allele specific oligonucleotide primers are useful in determining whether a subject is at risk of having breast or ovarian cancer, and also useful for characterizing a tumor. Primers direct amplification of a target polynucleotide prior to sequencing. These unique BRCA2 oligonucleotide primers for exons 2-27 shown in TABLE II were designed and produced specifically to optimize amplification of portions of BRCA2 which are to be sequenced.

The primers used to carry out this invention embrace oligonucleotides of sufficient length and appropriate sequence to provide initiation of polymerization. Environmental conditions conducive to synthesis include the presence of nucleoside triphosphates and an agent for polymerization, such as DNA polymerase, and a suitable temperature and pH. The primer is preferably single stranded for maximum efficiency in amplification, but may be double stranded. If double stranded, the primer is first treated to separate its strands before being used to prepare extension products. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the inducing agent for polymerization. The exact length of primer will depend on many factors, including temperature, buffer, and nucleotide composition. The oligonucleotide primer typically contains 18-28 bp plus in some cases an M13 "tail" for convenience.

Primers used to carry out this invention are designed to be substantially complementary to each strand of the genomic locus to be amplified. This means that the primers must be sufficiently complementary to hybridize with their respective strands under conditions which allow the agent for polymerization to perform. In other words, the primers should have sufficient complementarity with the 5' and 3' sequences flanking the mutation to hybridize therewith and permit amplification of the genomic locus.

Oligonucleotide primers of the invention are employed in the amplification process which is an enzymatic chain reaction that produces exponential quantities of polymorphic locus relative to the number of reaction steps involved. Typically, one primer is complementary to the negative (-) strand of the polymorphic locus and the other is complementary to the positive (+) strand. Annealing the primers to denatured nucleic acid followed by extension with an enzyme, such as the large

fragment of DNA polymerase I (Klenow) and nucleotides, results in newly synthesized + and - strands containing the target polymorphic locus sequence. Because these newly synthesized sequences are also templates, repeated cycles of denaturing, primer annealing, and extension results in exponential production of the region (*i.e.*, the target polymorphic locus sequence) defined by the primers. The product of the chain reaction is a discreet nucleic acid duplex with termini corresponding to the ends of the specific primers employed.

The oligonucleotide primers of the invention may be prepared using any suitable method, such as conventional phosphotriester and phosphodiester methods or automated embodiments thereof. In one such automated embodiment, diethylphosphoramidites are used as starting materials and may be synthesized as described by Beaucage, et al., Tetrahedron Letters, 22:1859-1862, 1981. One method for synthesizing oligonucleotides on a modified solid support is described in U.S. Patent No. 4,458,066.

The agent for polymerization may be any compound or system which will function to accomplish the synthesis of primer extension products, including enzymes. Suitable enzymes for this purpose include, for example, *E. coli* DNA polymerase I, Klenow fragment of *E. coli* DNA polymerase, polymerase muteins, reverse transcriptase, other enzymes, including heat-stable enzymes (*i.e.*, those enzymes which perform primer extension after being subjected to temperatures sufficiently elevated to cause denaturation), such as *Taq* polymerase. Suitable enzymes will facilitate combination of the nucleotides in the proper manner to form the primer extension products which are complementary to each polymorphic locus nucleic acid strand. Generally, the synthesis will be initiated at the 3' end of each primer and proceed in the 5' direction along the template strand, until synthesis terminates, producing molecules of different lengths.

The newly synthesized strand and its complementary nucleic acid strand will form a double-stranded molecule under hybridizing conditions described above and this hybrid is used in subsequent steps of the process. In the next step, the newly synthesized double-stranded molecule is subjected to denaturing conditions using any of the procedures described above to provide single-stranded molecules.

The steps of denaturing, annealing, and extension product synthesis can be repeated as often as needed to amplify the target polymorphic locus nucleic acid

sequence to the extent necessary for detection. The amount of the specific nucleic acid sequence produced will accumulate in an exponential fashion. Amplification is described in *PCR*. *A Practical Approach*, ILR Press, Eds. M. J. McPherson, P. Quirke, and G. R. Taylor, 1992.

The amplification products may be detected by Southern blots analysis, without using radioactive probes. In such a process, for example, a small sample of DNA containing a very low level of the nucleic acid sequence of the polymorphic locus is amplified, and analyzed via a Southern blotting technique or similarly, using dot blot analysis. The use of non-radioactive probes or labels is facilitated by the high level of the amplified signal. Alternatively, probes used to detect the amplified products can be directly or indirectly detectably labeled, for example, with a radioisotope, a fluorescent compound, a bioluminescent compound, a chemiluminescent compound, a metal chelator or an enzyme. Those of ordinary skill in the art will know of other suitable labels for binding to the probe, or will be able to ascertain such, using routine experimentation.

Sequences amplified by the methods of the invention can be further evaluated, detected, cloned, sequenced, and the like, either in solution or after binding to a solid support, by any method usually applied to the detection of a specific DNA sequence such as PCR, oligomer restriction (Saiki, et.al., Bio/Technology, 3:1008-1012, 1985), allele-specific oligonucleotide (ASO) probe analysis (Conner, et al., Proc. Natl. Acad. Sci. U.S.A., 80:278, 1983), oligonucleotide ligation assays (OLAs) (Landgren, et al., Science, 241:1007, 1988), and the like. Molecular techniques for DNA analysis have been reviewed (Landgren, et al., Science, 242:229-237, 1988).

Preferably, the method of amplifying is by PCR, as described herein and as is commonly used by those of ordinary skill in the art. Alternative methods of amplification have been described and can also be employed as long as the BRCA2 locus amplified by PCR using primers of the invention is similarly amplified by the alternative means. Such alternative amplification systems include but are not limited to self-sustained sequence replication, which begins with a short sequence of RNA of interest and a T7 promoter. Reverse transcriptase copies the RNA into cDNA and degrades the RNA, followed by reverse transcriptase polymerizing a second strand of DNA. Another nucleic acid amplification technique

is nucleic acid sequence-based amplification (NASBA) which uses reverse transcription and T7 RNA polymerase and incorporates two primers to target its cycling scheme. NASBA can begin with either DNA or RNA and finish with either, and amplifies to 10<sup>8</sup> copies within 60 to 90 minutes. Alternatively, nucleic acid can be amplified by ligation activated transcription (LAT). LAT works from a singlestranded template with a single primer that is partially single-stranded and partially double-stranded. Amplification is initiated by ligating a cDNA to the promoter oligonucleotide and within a few hours, and amplification is  $10^8\,$  to  $10^9\,$  fold. Another amplification system useful in the method of the invention is the  $\ensuremath{\text{\textbf{Q}}\beta}$ Replicase System. The Qβ replicase system can be utilized by attaching an RNA sequence called MDV-1 to RNA complementary to a DNA sequence of interest. Upon mixing with a sample, the hybrid RNA finds its complement among the specimen's mRNAs and binds, activating the replicase to copy the tag-along sequence of interest. Another nucleic acid amplification technique, ligase chain reaction (LCR), works by using two differently labeled halves of a sequence of interest which are covalently bonded by ligase in the presence of the contiguous sequence in a sample, forming a new target. The repair chain reaction (RCR) nucleic acid amplification technique uses two complementary and target-specific oligonucleotide probe pairs, thermostable polymerase and ligase, and DNA nucleotides to geometrically amplify targeted sequences. A 2-base gap separates the oligonucleotide probe pairs, and the RCR fills and joins the gap, mimicking normal DNA repair. Nucleic acid amplification by strand displacement activation (SDA) utilizes a short primer containing a recognition site for hincll with short overhang on the 5' end which binds to target DNA. A DNA polymerase fills in the part of the primer opposite the overhang with sulfur-containing adenine analogs. HincII is added but only cuts the unmodified DNA strand. A DNA polymerase that lacks 5' exonuclease activity enters at the site of the nick and begins to polymerize, displacing the initial primer strand downstream and building a new one which serves as more primer. SDA produces greater than 10<sup>7</sup>-fold amplification in 2 hours at 37°C. Unlike PCR and LCR, SDA does not require instrumented Temperature cycling.

Another method is a process for amplifying nucleic acid sequences from a DNA or RNA template which may be purified or may exist in a mixture of nucleic acids. The resulting nucleic acid sequences may be exact copies of the template, or may be modified. The process has advantages over PCR in that it increases the fidelity of copying a specific nucleic acid sequence, and it allows one to more efficiently detect a particular point mutation in a single assay. A target nucleic acid is amplified enzymatically while avoiding strand displacement. Three primers are used. A first primer is complementary to the first end of the target. A second primer is complementary to the second end of the target. A third primer which is similar to the first end of the target and which is substantially complementary to at least a portion of the first primer such that when the third primer is hybridized to the first primer, the position of the third primer complementary to the base at the 5' end of the first primer contains a modification which substantially avoids strand displacement. This method is detailed in U.S. Patent 5,593,840 to Bhatnagar et al. 1997, incorporated herein by reference.

Finally, recent application of DNA chips or microarray technology where DNA or oligonucleotides are immobilized on small solid support may also be used to rapidly sequence sample BRCA2 gene and analyze its expression. Typically, high density arrays of DNA fragment are fabricated on glass or nylon substrates by *in situ* light-directed combinatorial synthesis or by conventional synthesis followed by immobilization (Fodor *et al.* U.S. patent No. 5,445,934). Sample DNA or RNA may be amplified by PCR, labeled with a fluorescent tag, and hybridized to the microarray. Examples of this technology are provided in U.S. Patents 5,510, 270, U.S. 5,547,839, incorporated herein by reference.

All exonic and adjacent intronic sequences of the BRCA2 gene were obtained by end to end sequencing of five normal subjects in the manner described above followed by analysis of the data obtained. The data obtained provided us with the opportunity to establish the correct intronic/exonic structure of the BRCA2 gene. In addition, we evaluated six previously published normal polymorphisms (1342, 2457, 3199, 3624, 4035, and 7470) for correctness and frequency in the population, and to identify four additional polymorphisms not previously characterized (1093, 1593, 2908, and 9079).

### **GENE THERAPY**

The polynucleotide(s) which result from either sense or antisense transcription of any exon or the entire coding sequence or fragments of BRCA2 gene may be used for gene therapy. A variety of methods are known for gene transfer, any of which might be available for use.

Direct injection of Recombinant DNA in vivo:

- 1. Direct injection of "naked" DNA directly with a syringe and needle into a specific tissue, infused through a vascular bed, or transferred through a catheter into endothelial cells.
- 2. Direct injection of DNA that is contained in artificially generated lipid vesicles or other encapsulating vehicles.
- 3. Direct injection of DNA conjugated to a target receptor structure, such as a diptheria toxin, an antibody or other suitable receptor.
- 4. Direct injection by particle bombardment. For example, the DNA may be coated onto gold particles and shot into the cells.

### Human Artificial Chromosomes

The gene delivery approach involves the use of human chromosomes that have been stripped down to contain only the essential components for replication and the genes desired for transfer.

### Receptor-Mediated Gene Transfer

DNA is linked to a targeting molecule that will bind to specific cell-surface receptors, inducing endocytosis and transfer of the DNA into mammalian cells. One such technique uses poly-L-lysine to link asialoglycoprotein to DNA. An adenovirus is also added to the complex to disrupt the lysosomes and thus allow the DNA to avoid degradation and move to the nucleus. Infusion of these particles intravenously has resulted in gene transfer into hepatocytes.

### RECOMBINANT VIRUS VECTORS

Several vectors may be used in gene therapy. Among them are the Moloney Murine Leukemia Virus (MoMLV) Vectors, the adenovirus vectors, the Adeno-Associated Virus (AAV) vectors, the herpes simplex virus (HSV) vectors, the poxvirus vectors, the retrovirus vectors, and human immunodeficiency virus (HIV) vectors.

### GENE REPLACEMENT AND REPAIR

The ideal genetic manipulation for treatment of a genetic disease would be the actual replacement of the defective gene with a normal copy of the gene. Homologous recombination is the term used for switching out a section of DNA and replacing it with a new piece. By this technique, the defective gene may be replaced with a normal gene which expresses a functioning BRCA2 tumor growth inhibitor protein.

A complete description of gene therapy can also be found in "Gene Therapy A Primer For Physicians" 2d Ed. by Kenneth W. Culver, M.D. Publ. Mary Ann Liebert Inc. (1996). Two Gene Therapy Protocols for BRCA1 gene have been approved by the Recombinant DNA Advisory Committee for Jeffrey T. Holt et al. They are listed as 9602-148, and 9603-149 and are available from the NIH. Protocols for BRCA2 gene therapy may be similarly employed. The isolated BRCA2 gene may be synthesized or constructed from amplification products and inserted into a vector such as the LXSN vector.

### A BRCA2 POLYPEPTIDE OR ITS FUNCTIONAL EQUIVALENT

The growth of breast and ovarian cancer may be arrested or prevented by directly increasing the BRCA2 protein level where inadequate functional BRCA2 activity is responsible for breast and ovarian cancer. The cDNA and amino acid sequences of five novel BRCA2 haplotypes are disclosed herein (SEQ ID No:4-13). All or a fragment of BRCA2 protein may be used in therapeutic or prophylactic treatment of breast and ovarian cancer. Such a fragment may have a similar biological function as the native BRCA2 protein or may have a desired biological function as specified below. BRCA2 polypeptides or their functional equivalents including homologous and modified polypeptide sequences are also within the scope of the present invention. Changes in the native sequence may be advantageous in producing or using the BRCA2 derived polypeptides or functional equivalents suitable for therapeutic or prophylactic treatment of breast and ovarian cancer. For example, these changes may be desirable for producing resistance against *in vivo* proteolytic cleavage, for facilitating transportation and delivery of

therapeutic reagents, for localizing and compartmentalizing tumor suppressing agents, or for expression, isolating and purifying the target species.

There are a variety of methods to produce an active BRCA2 polypeptide or a functional equivalent as a tumor growth inhibitor. For example, one or more amino acids may be substituted, deleted, or inserted using methods well known in the art (Maniatis et al., 1982). Considerations of polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphiphathic nature of the amino acids play an important role in designing homologous polypeptide changes suitable for the intended treatment. In particular, conservative amino acid substitution using amino acids that are related in side-chain structure and charge may be employed to preserve the chemical and biological property. A homologous polyeptide typically contains at least 70% homology to the native sequence. Unnatural forms of the polypeptide may also be incorporated so long as the modification retains substantial biological activity. These unnatural polypeptides typically include structural mimics and chemical medications, which have similar three-dimensional structures as the active regions of the native BRCA2 protein. For example, these modifications may include terminal D-amino acids, cyclic peptides, unnatural amino acids side chains, pseudopeptide bonds, N-terminal acetylation, glycosylation, and biotinylation, etc. These unnatural forms of polypeptide may have a desired biological function, for example, they may be particularly robust in the presence of cellular or serum proteases and exopeptidase. An effective BRCA2 polypeptide or a functional equivalent may also be recognized by the reduction of the native BRCA2 protein. Regions of the BRCA2 protein may be systematically deleted to identify which regions are essential for tumor growth inhibitor activity. These smaller fragments of BRCA2 protein may then be subjected to structural and functional modification to derive therapeutically or prophylactically effective regiments. Finally, drugs, natural products or small molecules may be screened or synthesized to mimic the function of the BRCA2 protein. Typically, the active species retain the essential threedimensional shape and chemical reactivity, and therefore retain the desired aspects of the biological activity of the native BRCA2 protein. The activity and function of BRCA2 may include transactivation, granin, DNA repair among others. Functions of BRCA2 protein are also reviewed in Bertwistle and Ashworth, Curr. Opin. Genet. Dev. 8(1): 14-20 (1998) and Zhang et al., Cell 92:433-436 (1998). It will be

apparent to one skilled in the art that a BRCA2 polypeptide or a functional equivalent may be selected because such polypeptide or functional equivalent possesses similar biological activity as the native BRCA2 protein.

### EXPRESSION OF THE BRCA2 PROTEIN AND POLYPEPTIDE IN HOST CELLS

All or fragments of the BRCA2 protein and polypeptide may be produced by host cells that are capable of directing the replication and the expression of foreign genes. Suitable host cells include prokaryotes, yeast cells, or higher eukaryotic cells, which contain an expression vector comprising all or a fragment of the BRCA2 cDNA sequence (SEQ. ID No: 4, 6, 8, 10, or 12) operatively linked to one or more regulatory sequences to produce the intended BRCA2 protein or polypeptide. Prokaryotes may include gram negative or gram positive organisms, for example *E. coli* or *Bacillus* strains. Suitable eukaryotic host cells may include yeast, virus, and mamalian systems. For example, Sf9 insect cells and human cell lines, such as COS, MCF7, HeLa, 293T, HBL100, SW480, and HCT116 cells.

A broad variety of suitable expression vectors are available in the art. An expression vector typically contains an origin of replication, a promoter, a phenotypic selection gene (antibiotic resistance or autotrophic requirement), and a DNA sequence coding for all or fragments of the BRCA2 protein. The expression vectors may also include other operatively linked regulatory DNA sequences known in the art, for example, stability leader sequences, secretory leader sequences, restriction enzyme cleavage sequences, polyadenylation sequences, and termination sequences, among others. The essential and regulatory elements of the expression vector must be compatible with the intended host cell. Suitable expression vectors containing the desired coding and control regions may be constructed using standard recombinant DNA techniques known in the art, many of which are described in Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1989). For example, suitable origins of replication may include Col E1, SV4O viral and M13 origins of replication. Suitable promoters may be constitutive or inducible, for example, tac promoter, lac Z promoter, SV40 promoter, MMTV promoter, and LXSN promoter. Examples of selectable markers include neomycin, ampicillin, and hygromycin resistance and the like. Many suitable prokaryotic, viral and mammalian expression vectors may be obtained commercially, for example, from Invitrogen Corp., San Diego, CA or from Clontech, Palo Alto, CA. It may be desirable that the BRCA2 protein or polypeptide is produced as a fusion protein to enhance the expression in selected host cells, to detect the expression in transfected cells, or to simplify the purification process. Suitable fusion partners for the BRCA2 protein or polypeptide are well known in the art and may include  $\beta$ -galactosidase, glutathione-S-transferase, and poly-histidine tag.

Expression vectors may be introduced into host cells by various methods known in the art. The transformation procedure used depends upon the host to be transformed. Methods for introduction of vectors into host cells may include calcium phosphate precipitation, electrosporation, dextran-mediated transfection, liposome encapsulation, nucleus microinjection, and viral or phage infection, among others.

Once an expression vector has been introduced into a suitable host cell, the host cell may be cultured under conditions permitting expression of large amounts of the BRCA2 protein or polypeptide. The expression product may be identified by many approaches well known in the art, for example, sequencing after PCR-based amplification, hybridization using probes complementary to the desired DNA sequence, the presence or absence of marker gene functions such as enzyme activity or antibiotic resistance, the level of mRNA production encoding the intended sequence, immunological detection of a gene product using monoclonal and polyclonal antibodies, such as Western blotting or ELISA. The BRCA2 protein or polypeptides produced in this manner may then be isolated following cell lysis and purified using various protein purification techniques known in the art, for example, ion exchange chromatography, gel filtration chromatography and immunoaffinity chromatography.

It is generally preferred that whenever possible, longer fragments of BRCA2 protein or polypeptide are used, particularly to include the desired functional domains of BRCA2 protein. Expression of shorter fragments of DNA may be useful in generating BRCA2 derived immunogen for the production of anti-BRCA2 antibodies. It should, of course, be understood that not all expression vectors, DNA regulatory sequences or host cells will function equally well to express the BRCA2 protein or polypeptides of the present invention. However, one of ordinary skill in the art may make a selection among expression vectors, DNA regulatory

sequences, host cells, and codon usage in order to optimize expression using known technology in the art without undue experimentation. Studies of BRCA2 protein function and examples of genetic manipulation of BRCA2 protein are summarized in two recent review articles, Bertwistle and Ashworth, *Curr. Opin. Genet. Dev.* 8(1): 14-20 (1998) and Zhang *et al.*, *Cell* 92:433-436 (1998).

### IN VITRO SYNTHESIS AND CHEMICAL SYNTHESIS

Although it is preferred that fragments of the BRCA2 protein or polypeptides be obtained by overexpression in prokaryotic or eukaryotic host cells, the BRCA2 polypeptides or their functional equivalents may also be obtained by *in vitro* translation or synthetic means by methods known to those of ordinary skill in the art. For example, *in vitro* translation may employ an mRNA encoded by a DNA sequence coding for fragments of the BRCA2 protein or polypeptides. Chemical synthesis methodology such as solid phase synthesis may be used to synthesize a BRCA2 polypeptide structural mimic and chemically modified analogs thereof. The polypeptides or the modifications and mimic thereof produced in this manner may then be isolated and purified using various purification techniques, such as chromatographic procedures including ion exchange chromatography, gel filtration chromatography and immunoaffinity chromatography.

### PROTEIN REPLACEMENT THERAPY

The tumor suppressing function of BRCA2 suggests that various BRCA2 protein targeted therapies may be utilized in treating and preventing tumors in breast and ovarian cancer. The present invention therefore includes therapeutic and prophylactic treatment of breast and ovarian cancer using therapeutic pharmaceutical compositions containing the BRCA2 protein, polypeptides, or their functional equivalents. For example, protein replacement therapy may involve directly administering the BRCA2 protein, a BRCA2 polypeptide, or a functional equivalent in a pharmaceutically effective carrier. Alternatively, protein replacement therapy may utilize tumor antigen specific antibody fused to fragments of the BRCA2 protein, a polypeptide, or a functional equivalent to deliver anti-cancer regiments specifically to the tumor cells.

To prepare the pharmaceutical compositions of the present invention, an active BRCA2 protein, a BRCA2 polypeptide, or its functional equivalent is combined with a pharmaceutical carrier selected and prepared according to conventional pharmaceutical compounding techniques. A suitable amount of the composition may be administered locally to the site of a tumor or systemically to arrest the proliferation of tumor cells. The methods for administration, may include parenteral, oral, or intravenous, among others according to established protocols in the art.

Pharmaceutically acceptable solid or liquid carriers or components which may be added to enhance or stabilize the composition, or to facilitate preparation of the composition include, without limitation, syrup, water, isotonic solution, 5 % glucose in water or buffered sodium or ammonium acetate solution, oils, glycerin, alcohols, flavoring agents, preservatives, coloring agents, starches, sugars, diluents, granulating agents, lubricants, binders, and sustained release materials. The dosage at which the therapeutic compositions are administered may vary within a wide range and depends on various factors, such as the stage of cancer progression, the age and condition of the patient, and may be individually adjusted.

### DIAGNOSTIC REAGENTS

The BRCA2 protein, polypeptides, their functional equivalents, antibodies, and polynucleotides may be used in a wide variety of ways in addition to gene therapy and protein replacement therapy. They may be useful as diagnostic reagents to measure normal or abnormal activity of BRCA2 at the DNA, RNA, and protein level. The present invention therefore encompasses the diagnostic reagents derived from the BRCA2 cDNA and protein sequences as set forth in SEQ. ID. Nos: 4-13. These reagents may be utilized in methods for monitoring disease progression, for determining patients suited for gene and protein replacement therapy, or for detecting the presence or quantifying the amount of a tumor growth inhibitor following such therapy. Such methods may involve conventional histochemical techniques, such as obtaining a tumor tissue from the patient, preparing an extract and testing this extract for tumor growth or metabolism. For example, the test for tumor growth may involve measuring abnormal BRCA2 activity using conventional diagnostic assays, such as Southern, Northern, and Western blotting, PCR, RT-PCR, and immunoprecipitation. In

biopsies of tumor tissues, the loss of BRCA2 expression in tumor tissue may be verified by RT-PCR and Northern blotting at the RNA level. A Southern blot analysis, genomic PCR, or fluorescence in situ hybridization (FISH) may also be performed to examine the mutations of BRCA2 at the DNA level. And, a Western blotting, protein truncation assay, or immunoprecipitation may be utilized to analysis the effect at the protein level.

These diagnostic reagents are typically either covalently or non convalently attached to a detectable label. Such a label includes a radioactive label, a colorimetric enzyme label, a fluorescence label, or an epitope label. Frequently, a reporter gene downstream of the regulatory sequences is fused with the BRCA2 protein or polypeptide to facilitate the detection and purification of the target species. Commonly used reporter genes in BRCA2 fusion proteins include  $\beta$ -galactosidase and luciferase gene.

The BRCA2 protein, polypeptides, their functional equivalents, antibodies, and polynucleotides may also be useful in the study of the characteristics of BRCA2 proteins, such as structure and function of BRCA2 in oncogenesis or subcellular localization of BRCA2 protein in normal and cancerous cell. For example, yeast two-hybrid system has been used in the study of cellular function of BRCA2 to identify the regulator and effector of BRCA2 tumor suppressing function (Sharan et al., Nature 386:804-810 (1997) and Katagiri et al., Genes, Chromosomes & Cancer 21:217-222 (1988)). In addition, the BRCA2 protein, polypeptides, their functional equivalents, antibodies, and polynucleotides may also be used in *in vivo* cell based and *in vitro* cell free assays to screen natural products and synthetic compounds which may mimic, regulate or stimulate BRCA2 protein function.

### ANTISENSE INHIBITION

Antisense suppression of endogenous BRCA2 expression may assess the effect of BRCA2 protein on cell growth inhibition using known method in the art (Crooke, *Annu. Rev. Pharmacol. Toxicol.* 32:329-376 (1992) and Robinson-Benion and Holt, *Methods Enzymol.* 254:363-375 (1995)). Given the cDNA sequence as set forth in SEQ ID. NO: 4, 6, 8, 10, and 12, one of skill in the art can readily obtain anti-sense strand of DNA and RNA sequences to interfere with the production of wild-type BRCA2 protein or the mutated form of BRCA2 protein. Alternatively,

antisense oligonucleotide may be designed to target the control sequences of BRCA2 gene to reduce or prevent the expression of the endogenous BRCA2 gene.

### **ANTIBODIES**

The BRCA2 protein, polypeptides, or their functional equivalents may be used as immunogens to prepare polyclonal or monoclonal antibodies capable of binding the BRCA2 derived antigens in a known manner (Harlow & Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1988). These antibodies may be used for the detection of the BRCA2 protein, polypeptides, or a functional equivalent in an immunoassay, such as ELISA, Western blot, radioimmunoassay, enzyme immunoassay, and immunocytochemistry. Typically, an anti-BRCA2 antibody is in solution or is attached to a solid surface such as a plate, a particle, a bead, or a tube. The antibody is allowed to contact a biological sample or a blot suspected of containing the BRCA2 protein or polypeptide to form a primary immunocomplex. After sufficient incubation period, the primary immunocomplex is washed to remove any non-specifically bound species. The amount of specifically bound BRCA2 protein or polypeptide may be determined using the detection of an attached label or a marker, such as a radioactive, a fluorescent, or an enzymatic label. Alternatively, the detection of BRCA2 derived antigen is allowed by forming a secondary immunocomplex using a second antibody which is attached with a such label or marker. The antibodies may also be used in affinity chromatography for isolating or purifying the BRCA2 protein, polypeptides or their functional equivalents.

### **EXAMPLE 1**

### <u>Determination of the Coding Sequence Haplotypes of the BRCA2 Gene From Normal Individuals</u>

Approximately 150 volunteers were screened in order to identify individuals with no cancer history in their immediate family (i.e. first and second degree relatives). Each person was asked to fill out a hereditary cancer prescreening questionnaire (See TABLE I). Five of these were randomly chosen for end-to-end sequencing of their BRCA2 gene. A first degree relative is a parent, sibling, or

offspring. A second degree relative is an aunt, uncle, grandparent, grandchild, niece, nephew, or half-sibling.

Genomic DNA was isolated from white blood cells of five normal subjects selected from analysis of their answers to the questions above. Dideoxy sequence analysis was performed following polymerase chain reaction amplification.

All exons of the BRCA2 gene were subjected to direct dideoxy sequence analysis by asymmetric amplification using the polymerase chain reaction (PCR) to generate a single stranded product amplified from this DNA sample. Shuldiner, *et al.*, *Handbook of Techniques in Endocrine Research*, p. 457-486, DePablo, F., Scanes, C., eds., Academic Press, Inc., 1993. Fluorescent dye was attached for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer<sup>®</sup> cat# 401628). DNA sequencing was performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) automated sequencer (Model 377). The software used for analysis of the resulting data was "Sequence Navigator" purchased through ABI.

### 1. Polymerase Chain Reaction (PCR) Amplification

Genomic DNA (100 nanograms) extracted from white blood cells of five normal subjects. Each of the five samples was sequenced end to end. Each sample was amplified in a final volume of 25 microliters containing 1 microliter (100 nanograms) genomic DNA, 2.5 microliters 10X PCR buffer (100 mM Tris, pH 8.3, 500 mM KCl, 1.2 mM MgCl<sub>2</sub>), 2.5 microliters 10X dNTP mix (2 mM each nucleotide), 2.5 microliters forward primer, 2.5 microliters reverse primer, and 1 microliter Taq polymerase (5 units), and 13 microliters of water.

The primers in TABLE II below were used to carry out amplification of the various sections of the BRCA2 gene samples. The primers were synthesized on an DNA/RNA Synthesizer Model  $394^{\$}$ .

Thirty-five cycles were performed, each consisting of denaturing (95°C; 30 seconds), annealing (55°C; 1 minute), and extension (72°C; 90 seconds), except during the first cycle in which the denaturing time was increased to 5 minutes, and during the last cycle in which the extension time was increased to 5 minutes.

PCR products were purified using Qia-quick<sup>®</sup> PCR purification kits (Qiagen<sup>®</sup>, cat# 28104; Chatsworth, CA). Yield and purity of the PCR product are determined spectrophotometrically at  $OD_{260}$  on a Beckman DU 650 spectrophotometer.

### 2. Dideoxy Sequence Analysis

Fluorescent dye was attached to PCR products for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer<sup>®</sup> cat # 401628). DNA sequencing was performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) Foster City, CA., automated sequencer (Model 377). The software used for analysis of the resulting data was "Sequence Navigator<sup>®</sup>" purchased through ABI.

### 3. RESULTS

Based upon the sequencing of the five normal individuals, it was determined that the standard sequence found in both GenBank and BIC were inaccurate. In Genbank, a 10 bp stretch (5'-TTTATTTTAG-3') was mistakenly listed as exonic at the 5' end of exon 5 while it should be intronic which would not be included in the cDNA and resultant protein. In addition, a more detrimental error that has the significant potential to lead to an incorrect diagnosis of breast cancer propensity exists in both Genbank and BIC: a sequence of 16 bp (5'-GTGTTCTCATAAACAG-3') should be at the end of exon 15, but instead is listed at the beginning of exon 16 in the database. The disclosure and listing of GenBank is shown in Figure 1. The correct intron/exon sequence of BRCA2 is presented in Figure 2, wherein,

- (1) a 10 bp stretch (5'-TTTATTTTAG-3') is intronic at 3' end of intron 4, rather than at the 5' end of exon 5 (corrected exon 5 is listed as SEQ. ID. NO: 1) and
- (2) a 16 bp stretch (5'-GTGTTCTCATAAACAG-3') is exonic at the 3' end of exon 15, rather than at the 5' end of exon 16 (corrected exons 15 and 16 are listed as SEQ. ID. No: 2 and 3 respectively)

The BIC BRCA2 sequence also contains sequence errors in which a strech of nine nucleotides at positions 5554-5460 is listed as CGTTTGTGT (amino acids: Arg-

Leu-Cys). The correct sequence at these positions is GTTTGTGTT (amino acids: Val-Cys-Val). In addition, the BIC BRCA2 nuclotides at positions 2024 (codon 599), 4553 (codon 1442), 4815 (codon 1529), 5841 (codon 1871), and 5972 (codon 1915) are T, T, A, C, and T respectively, wherein the correct nucleotides at these positions are C, C, G, T, and C respectively. Among them, the nuclotide errors at codon 599, 1442, 1915 result in amino acids changes.

Additional differences in the nucleic acids of the five normal individuals were found in ten polymorphic locations. The changes and their positions are found in TABLE III. The individual haplotypes of each chromosome of BRCA2 are displayed in FIGURE 3. In each case, the initial haplotype reported in Genbank (accession number U43746) was subtracted to determine the new haplotypes OMI 1-5. Thus, the Genbank sequence only represents 50% of the haplotypes found; the five new BRCA2 (orm 1-5) DNA sequences are shown as SEQ. ID. NO: 4, 6, 8, 10, and 12, respectively (See FIGURE 3), and the corresponding polypeptides are listed as SEQ. ID. NO: 5, 7, 9, 11, and 13 respectively. In combination, these seven haplotypes represent a functional allele profile for the BRCA2 gene.

The data show that for each of the samples, all exons of BRCA2 were identical except in the region of ten polymorphisms. Six of these polymorphisms were previously identified (Tartigan et al., Nature Genetics 12: 333-337 (1996); Phelan et al., Nature Genetics 13: 120-122 (1996); Couch et al., Nature Genetics 13: 123-125 (1996); Teng, et al., Nature Genetics 13: 241-244 (1996); Schubert et al. 60: 1031-1040 (1997)), but four were unique to this work. Even though the individual polymorphisms may have been identified, none of these complete haplotypes has been previously determined.

### **TABLE I**

### Hereditary Cancer Pre-Screening Questionnaire

### Part A: Answer the following questions about your family

- 1. To your knowledge, has anyone in your family been diagnosed with a very specific hereditary colon disease called Familial Adenomatous Polyposis (FAP)?
- 2. To your knowledge, have you or any aunt had breast cancer diagnosed before the age 35?
- 3. Have you had Inflammatory Bowel Disease, also called Crohn's Disease or Ulcerative Colitis, for more than 7 years?

### Part B: Refer to the list of cancers below for your responses only to questions in Part B

Bladder Cancer Lung Cancer Pancreatic Cancer
Breast Cancer Gastric Cancer Prostate Cancer
Colon Cancer Malignant Melanoma Renal Cancer
Endometrial Cancer Ovarian Cancer Thyroid Cancer

- 4. Have your mother or father, your sisters or brothers or your children had any of the listed cancers?
- Have there been diagnosed in your <u>mother</u>'s brothers or sisters, or your <u>mother</u>'s parents more than one of the cancers in the above list?
- 6. Have there been diagnosed in your <u>father</u>'s brothers or sisters, or your <u>father</u>'s parents <u>more</u> than <u>one</u> of the cancers in the above list?

### Part C: Refer to the list of relatives below for responses only to questions in Part C

You Your mother

Your sisters or brothers Your mother's sisters or brothers (maternal aunts &

uncles)

Your children Your mother's parents (maternal grandparents)

- 7. Have there been diagnosed in these relatives <u>2 or more identical</u> types of cancer? Do not count "simple" skin cancer, also called basal cell or squamous cell skin cancer.
- 8. Is there a <u>total of 4 or more</u> of any cancers in the list of relatives above other than "simple" skin cancers?

### Part D: Refer to the list of relatives below for responses only to questions in Part D.

You Your father

Your sisters or brothers Your father's sisters or brothers (paternal aunts and

uncles)

Your children Your father's parents (paternal grandparents)

- 9. Have there been diagnosed in these relatives <u>2 or more identical</u> types of cancer? Do not count "simple" skin cancer, also called basal cell or squamous cell skin cancer.
- 10. Is there a total of 4 or more of any cancers in the list of relatives above other than "simple" skin cancers?

© Copyright 1996, OncorMed, Inc.

E C	-	SEQUENCE (5' TO 3') NOTE: M13 TAIL INCLUDED	Oligo	PCR	SEQ. ID.
<u> </u>		M13 FORWARD = TGT AAA ACG ACG GCC AGT M13 REVERSE = CAG GAA ACA GCT ATG ACC	Length	Product Length	Number
10C	BRCA2-10CF/M13F	5'-TGT AAA ACG ACG GCC AGT CAG CAT CTT GAA TCT CAT ACA G-3'	40	508	33
10C	BRCA2-10CRII	5'-AGA CAG AGG TAC CTG AAT C-3'	19		34
11	BRCA2-11AF-M13	5'- TGT AAA ACG ACG GCC AGT TGG TAC TTT AAT TTT GTC ACT T-3'	40	304	35
11	BRCA2-11AR-M13	5'-CAG GAA ACA GCT ATG ACC TGC AGG CAT GAC AGA GAA T-3'	37		36
1	BRCA2-11BF	5'-AAG AAG CAA AAT GTA ATA AGG A-3'	22	411	37
11	BRCA2-11BR	5'-CAT TTA AAG CAC ATA CAT CTT G-3'	22		38
17	BRCA2-11CF	5'-TCT AGA GGC AAA GAA TCA TAC-3'	21	349	39
1	BRCA2-11CR	5'-CAA GAT TAT TCC TTT CAT TAG C-3'	22		40
11	BRCA2-11DF	5'-AAC CAA AAC ACA AAT CTA AGA G-3'	22	344	41
11	BRCA2-11DR	5'-GTC ATT TTT ATA TGC TGC TTT AC-3'	23		42
1	BRCA2-11EF	5'-GGT TTT ATA TGG AGA CAC AGG-3'	21	369	43
11	BRCA2-11ER	5'-GTA TTT ACA ATT TCA ACA CAA GC-3'	23		44
1	BRCA2-11FF	5'-ATC ACA GTT TTG GAG GTA GC-3'	50	368	45
7	BRCA2-11FR	5'-CTG ACT TCC TGA TTC TTC TAA-3'	21		46
7-	BRCA2-11GF	5'-CTC AGA TGT TAT TTT CCA AGC-3'	21	366	47
17	BRCA2-11GR	5'-CTG TTA AAT AAC CAG AAG CAC-3'	21		48
11	BRCA2-11HF	5'-AGG TAG ACA GCA AGC-3'	18	360	49
1	BRCA2-11HR	5'-GTA ATA TCA GTT GGC ATT TAT T-3'	22		20
1-	BRCA2-11IF	5'-TGC AGA GGT ACA TCC AAT AAG-3'	21	326	51
1	BRCA2-11IR	5'-GAT CAG TAA ATA GCA AGT CCG-3'	21		52
17	BRCA2-11JF	5'-TAC TGA AAA TGA AGA TAA CAA AT-3'	23	477	53

—					ı –							Γ				Γ		I	<u> </u>		Ι	
SEQ. ID. Number	54	55	26	22	28	29	09	61	62	63	64	65	99	29	89	69	20	7.1	72	73	74	•
PCR Product Length		382		374		409		306		383		355		337		360		458		344		_
Oligo Length	75	35	35	22	19	20	22	35	35	22	20	20	20	21	21	20	20	35	37	22	21	-
SEQUENCE (5' TO 3') NOTE: M13 TAIL INCLUDED M13 FORWARD = TGT AAA ACG ACG GCC AGT M13 REVERSE = CAG GAA ACA GCT ATG ACC	5'-ATT TTG TTC TTT CTT ATG TCA G-3'	5'-TGT AAA ACG ACG GCC AGT CTA CTA AAA CGG AGC AA-3'	5'-CAG GAA ACA GCT ATG ACC GTA TGA AAA CCC AAC AG-3'	5'-CAC AAA ATA CTG AAA GAA AGT G-3'	5'-GGC ACC ACA GTC TCA ATA G-3'	5'-GCA AAG ACC CTA AAG TAC AG-3'	5'-CAT CAA ATA TTC CTT CTC TAA G-3'	5'-TGT AAA ACG ACG GCC AGT GAA AAT TCA GCC TTA GC-3'	5'- CAG GAA ACA GCT ATG ACC ATC AGA ATG GTA GGA AT-3'	5'-GTA CTA TAG CTG AAA ATG ACA A-3'	5'-ACC ACT GGC TAT CCT AAA TG-3'	5'-TGA AGA TAT TTG CGT TGA GG-3'	5'-GTC AGC AAA AAC CTT ATG TG-3'	5'-ACG AAA ATT ATG GCA GGT TGT-3'	5'-CTT GTC TTG CGT TTT GTA ATG-3'	5'-GCT TCA TAA GTC AGT CTC AT-3'	5'-TCA AAT TCC TCT AAC ACT CC-3'	5'-TGT AAA ACG ACG GCC AGT TAC AGC AAG TGG AAA GC-3'	5'-CAG GAA ACA GCT ATG ACC AAG TTT CAG TTT TAC CAA T-3'	5'-GTT CTT CAG AAA ATA ATC ACT C-3'	5'-TGT AAA AAG AGA ATG TGT GGC-3'	_
Label	BRCA2-11JR	BRCA2-11KF-M13	BRCA2-11KR-M13	BRCA2-11LF	BRCA2-11LR	BRCA2-11MF	BRCA2-11MR	BRCA2-11NF-M13	BRCA2-11NR-M13	BRCA2-110F	BRCA2-110R	BRCA2-11PF	BRCA2-11PR	BRCA2-11QF	BRCA2-11QR	BRCA2-11RF	BRCA2-11RR	BRCA2-11SF-M13	BRCA2-11SR-M13	BRCA2-11TF	BRCA2-11TR	_
Exon	11	11	=	=	=	=	=	=	=	11	11	11	7	11	11	11	11	11	=	7	=	

SEQ. ID. Number	22	9/	77	78	79	80	81	82	83	84	85	86	87	88	89	06	91	92	93	94	95	
PCR Product Length	328		391		310		391		284		394		282		275		355		296		304	
Oligo Length	39	39	42	40	21	28	22	22	21	38	24	20	19	20	39	38	41	39	38	39	39	
SEQUENCE (5' TO 3') NOTE: M13 TAIL INCLUDED M13 FORWARD = TGT AAA ACG ACG AGT M13 REVERSE = CAG GAA ACA GCT ATG ACC	5'-TGT AAA ACG ACG GCC AGT ACT TTT TCT GAT GTT CCT GTG-3'	5'-CAG GAA ACA GCT ATG ACC TAA AAA TAG TGA TTG GCA ACA-3'	5'-TGT AAA ACG ACG GCC AGT AGT GGT GTT TTA AAG TGG TCA AAA-3'	5'-CAG GAA ACA GCT ATG ACC GGA TCC ACC TGA GGT CAG AAT A-3'	5'-TAA CAT TTA AGC ATC CGT TAC-3'	5'-AAA CGA GAC TTT TCT CAT ACT GTA TTA G-3'	5'-ACC ATG TAG CAA ATG AGG GTC T-3'	5'-GCT TTT GTC TGT TTT CCT CCA A-3'	5'-CCA GGG GTT GTG CTT TTT AAA-3'	5'-CAG GAA ACA GCT ATG ACC ACT CTG TCA TAA AAG CCA TC-3'	5'-TTT GGT TTG TTA TAA TTG TTT TTA-3'	5'-CCA ACT TTT TAG TTC GAG AG-3'	5'-TTC AGT ATC ATC CTA TGT G-3'	5'-AGA AAC CTT AAC CCA TAC TG-3'	BRCA2-18FUT/M13-AF 5'-TGT AAA ACG ACG GCC AGT GAA TTC TAG AGT CAC ACT TCC-3'	5'-CAG GAA ACA GCT ATG ACC TTT AAC TGA ATC AAT GAC TG-3'	5'-TGT AAA ACG ACG GCC AGT AAG TGA ATA TTT TTA AGG CAG TT-3'	BRCA2-19FUT/M13-R 5'-CAG GAA ACA GCT ATG ACC AAG AGA CCG AAA CTC CAT CTC-3'	5'-TGT AAA ACG ACG GCC AGT CAC TGT GCC TGG CCT GAT AC-3'	5'-CAG GAA ACA GCT ATG ACC ATG TTA AAT TCA AAG TCT CTA-3'	5'-TGT AAA ACG ACG GCC AGT GGG TGT TTT ATG CTT GGT TCT-3'	
Label	BRCA2-11UF-M13	BRCA2-11UR-M13	BRCA2-12F/M13F	BRCA2-12R/M13R	BRCA2/13-2F	BRCA2/13-2R	BRCA2-14F	BRCA2-14AR	BRCA2-15-2F	BRCA2-15FUT/M13-R	BRCA2-16AF	BRCA2-16AR	BRCA2-17F	BRCA2-17AR	BRCA2-18FUT/M13-A	BRCA2-18R/M13R	BRCA2-19F/M13F	BRCA2-19FUT/M13-R	BRCA2-20F/M13F	BRCA2-20R/M13R	BRCA2-21F/M13F	
Exon	1	17	12	12	13	13	14	14	15	15	16	16	17	17	18	18	19	19	70	20	21	

Exon	Label	SEQUENCE (5' TO 3') NOTE: M13 TAIL INCLUDED M13 FORWARD = TGT AAA ACG ACG GCC AGT M13 REVERSE = CAG GAA ACA GCT ATG ACC	Oligo Length	PCR Product Length	SEQ. ID. Number
21	BRCA2-21R/M13R	5'-CAG GAA ACA GCT ATG ACC CAT TTC AAC ATA TTC CTT CCT G-3'	40		96
22	BRCA2-22F-1A	5'-AAC CAC ACC CTT AAG ATG A-3'	19	453	97
22	BRCA2-22R-1A	5'-GCA TTA GTA GTG GAT TTT GC-3'	20		86
23	BRCA2-23FII	5'-TCA CTT CCA TTG CAT C-3'	16	290	66
23	BRCA2-23RII	5'-TGC CAA CTG GTA GCT CC-3'	17		100
24	BRCA2-24 2F	5'-TAC AGT TAG CAG CGA CAA AA-3'	20	373	101
24	BRCA2-24R/M13R	5'-CAG GAA ACA GCT ATG ACC ATT TGC CAA CTG GTA GCT CC-3'	38		102
25	BRCA2-25F-7/23	5'-GCT TTC GCC AAA TTC AGC TA-3'	20	427	103
25	BRCA2-25R-7/23	5'-TAC CAA AAT GTG TGG TGA TG-3'	20		104
26	BRCA2/26-2F	5'-AAT CAC TGA TAC TGG TTT TG-3'	20	530	105
26	BRCA2/26-2R	5'-TAT ACT TAC AGG AGC CAC AT-3'	20		106
27A	BRCA2-27AF-1A	5'-CTG TGT GTA ATA TTT GCG-3'	18	495	107
27A	BRCA2-27AR/M13R	5'-CAG GAA ACA GCT ATG ACG GCA AGT TCT TCG TCA GCT ATT G-3'	40		108
278	BRCA2-27BF/M13F	5'-TGT AAA ACG ACG GCC AGT GAA TTC TCC TCA GAT GAC TCC A-3'	40	417	109
27B	BRCA2-27BR/M13R	5'-CAG GAA ACA GCT ATG ACC TCT TTG CTC ATT GTG CAA CA-3'	38		110

TABLE III NORMAL PANEL TYPING

Position nt/codon	Nucleotide Change	Amino Acid Change	-	2	8	4	5	Frequency
1093/289	AAT → CAT	Asn → His	A/A	A/C	A/A	A/A	A/C	A
1342/372	AAT → CAT	Asn → His	A/C	A/A	A/C	A/C	A/C	A = 0.6 C = 0.4
1593/455	TC <u>A</u> → TC <u>G</u>	Ser → Ser	A/A	A/A	A/A	A/A	A/G	A = 0.9 G = 0.1
2457/743	CA <u>I</u> → CA <u>C</u>	His → His	Т/Т	C/T	Т/Т	Т/Т	C/T	T = 0.8 C = 0.2
2908/894	GTA → <u>A</u> TA	Val → Ile	9/9	9/9	9/9	9/9	A/G	G = 0.9 A = 0.1
3199/991	<u>A</u> AC → <u>G</u> AC	Asn → Asp	A/A	A/G	A/A	A/A	A/G	A = 0.8 G = 0.2

TABLE III NORMAL PANEL TYPING

Position nt/codon	Nucleotide Change	Amino Acid Change	-	2	က	4	5	Frequency
3624/1132	AA <u>A</u> → AA <u>G</u>	Lys→Lys	A/A	A/G	A/A	A/G	A/A	A = 0.8 G = 0.2
4035/1269	GT <u>I</u> → GT <u>C</u>	Val → Val	С/Т	T/T	T/T	T/T	Т/Т	T = 0.9 C = 0.1
7470/2414	TC <u>A</u> → TC <u>G</u>	Ser → Ser	Α/A	A/G	A/A	A/G	A/A	A = 0.8 G = 0.2
9079/2951	GCC → ACC	Ala → Thr	9/9	9/9	9/9	9/9	A/G	G = 0.9 A = 0.1

# **EXAMPLE 2**

5

10

15

20

25

30

# <u>Determination Of A Normal Individual Using BRCA2<sup>(OMI 1-5)</sup> and The Ten</u> Polymorphisms For Reference

A person skilled in the art of genetic susceptibility testing will find the present invention useful for:

- a) identifying individuals having a normal BRCA2 gene;
- avoiding misinterpretation of normal polymorphisms found in the normal population.

Sequencing was carried out as in EXAMPLE 1 using a blood sample from the patient in question. However, the BRCA2 sequences were used for reference and any polymorphic sites seen in the patient were compared to the nucleic acid sequences listed above for normal codons at each polymorphic site. A normal sample is one which is comparable to the BRCA2 equences and contains only minor variations which occur at minor polymorphic sites. The allelic variations which occur at each of the polymorphic sites are paired here for reference.

- AAT (Asn) and CAT (His) at position 1093 (codon 289)
- CAT (His) and AAT (Asn) at position 1342 (codon 372)
- TCA (Ser) and TCG (Ser) at position 1593 (codon 455)
- CAT (His) and CAC (His) at position 2457 (codon 743)
- GTA (Val) and ATA (IIe) at position 2908 (codon 894)
- <u>A</u>AC (Asn) and <u>G</u>AC (Asp) at position 3199 (codon 991)
- AAA (Lys) and AAG (Lys) at position 3624 (codon 1132)
- GTT (Val) and GTC (Val) at position 4035 (codon 1269)
- TCA (Ser) and TCG (Ser) at position 7470 (codon 2414)
- GCC (Ala) and ACC (Thr) at position 9079 (codon 2951)

The availability of these polymorphic pairs provides added assurance that one skilled in the art can correctly interpret the polymorphic variations without mistaking a normal variation for a mutation.

All exons of the BRCA2 gene are subjected to direct dideoxy sequence analysis by asymmetric amplification using the polymerase chain reaction (PCR) to generate a single stranded product amplified from this DNA sample. Shuldiner, et

10

15

20

25

30

al., Handbook of Techniques in Endocrine Research, p. 457-486, DePablo, F., Scanes, C., eds., Academic Press, Inc., 1993. Fluorescent dye is attached for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer<sup>®</sup> cat# 401628). DNA sequencing is performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) automated sequencer (Model 377). The software used for analysis of the resulting data is "Sequence Navigator" purchased through ABI.

# 1. Polymerase Chain Reaction (PCR) Amplification

The PCR primers used to amplify a patient's sample BRCA2 gene are listed in TABLE II. The primers were synthesized on a DNA/RNA Synthesizer Model 394<sup>®</sup>. Thirty-five cycles are of amplification are performed, each consisting of denaturing (95°C; 30 seconds), annealing (55°C; 1 minute), and extension (72°C; 90 seconds), except during the first cycle in which the denaturing time is increased to 5 minutes and during the last cycle in which the extension time is increased to 5 minutes.

PCR products are purified using Qia-quick<sup>®</sup> PCR purification kits (Qiagen<sup>®</sup>, cat# 28104; Chatsworth, CA). Yield and purity of the PCR product are determined spectrophotometrically at OD<sub>260</sub> on a Beckman DU 650 spectrophotometer.

#### 2. Dideoxy Sequence Analysis

Fluorescent dye is attached to PCR products for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer<sup>®</sup> cat# 401628). DNA sequencing is performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) Foster City, CA., automated sequencer (Model 377). The software used for analysis of the resulting data is "Sequence Navigator<sup>®</sup>" purchased through ABI. The BRCA2<sup>(omi 1-5)</sup> sequences were entered sequentially into the Sequence Navigator software as the standards for comparison. The Sequence Navigator software compares the patient sample sequence to each BRCA2 <sup>(omi 1-5)</sup> standard, base by base. The Sequence Navigator highlights all differences between the standards (omi 1-5) and the patient's sample sequence.

A first technologist checks the computerized results by comparing visually the BRCA2 (OTH 1-5) standards against the patient's sample, and again highlights any differences between the standard and the sample. The first primary technologist then interprets the sequence variations at each position along the sequence. Chromatograms from each sequence variation are generated by the Sequence Navigator and printed on a color printer. The peaks are interpreted by the first primary technologist and a second primary technologist. A secondary technologist then reviews the chromatograms. The results are finally interpreted by a geneticist. In each instance, a variation is compared to known normal polymorphisms for position and base change.

#### 3. Results

5

10

15

20

25

30

The patient's BRCA2 sequence was found to be heterozygous at seven nucleotide positions: 1093 (A/C), 1342 (A/C), 1593 (A/G), 2457 (C/T), 2908 (A/G), 3199 (A/G) and 9079 (A/G). In addition, this changes five amino acids in the polypeptide product: Asn to His at codon 289, Asn to His at codon 372, Val to Ile at codon 894, Asn to Asp at codon 991, and Ala to Thr at codon 2951. The question arises whether any or all of these changes have significance to the patient. Comparison of the patient's results to the BRCA (omi 1-5) haplotypes demonstrates that it matches one of the BRCA2 omi standards (#5), and thus the patient sample is interpreted as carrying a normal gene sequence without causing any elevation in their risk status for breast cancer.

# **EXAMPLE 3**

# DETERMINING THE PRESENCE OF A MUTATION IN EXON 11 OF THE BRCA2 GENE USING BRCA2<sup>(omi1-5)</sup>

A person skilled in the art of genetic susceptibility testing will find the present invention useful for determining the presence of a known or previously unknown mutation in the BRCA2 gene. A list of mutations of BRCA2 is publicly available in the Breast Cancer Information Core at http://www.nchgr.nih.gov/dir/lab\_transfer/bic. This data site became publicly available on November 1, 1995. Friend, S. *et al. Nature Genetics* 11:238, (1995).

In this example, a mutation in exon 11 is characterized by amplifying the region of the mutation with a primer set which amplifies the region of the mutation. Sequencing was carried out as in Example 1 using a blood sample from the patient in question. Specifically, exon 11 of the BRCA2 gene is subjected to direct dideoxy sequence analysis by asymmetric amplification using the polymerase chain reaction (PCR) to generate a single stranded product amplified from this DNA sample. Shuldiner, et al., Handbook of Techniques in Endocrine Research, p. 457-486, DePablo, F., Scanes, C., eds., Academic Press, Inc., 1993. Fluorescent dye is attached for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer<sup>®</sup> cat# 401628). DNA sequencing is performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) automated sequencer (Model 377). The software used for analysis of the resulting data is "Sequence Navigator" purchased through ABI.

15

20

25

10

5

# 1. Polymerase Chain Reaction (PCR) Amplification

Genomic DNA (100 nanograms) extracted from white blood cells of the subject is amplified in a final volume of 25 microliters containing 1 microliter (100 nanograms) genomic DNA, 2.5 microliters 10X PCR buffer (100 mM Tris, pH 8.3, 500 mM KCl, 1.2 mM MgCl<sub>2</sub>), 2.5 microliters 10X dNTP mix (2 mM each nucleotide), 2.5 microliters forward primer (BRCA2-11Q-F, 10 micromolar solution), 2.5 microliters reverse primer (BRCA2-11Q-R, 10 micromolar solution), and 1 microliter Tag polymerase (5 units), and 13 microliters of water.

The PCR primers used to amplify segment Q of exon 11 (where the mutation 6174delT is found) are as follows:

BRCA2-11Q-F: 5'- ACG' AAA' ATT' ATG' GCA' GGT' TGT-3'

BRCA2-11Q-R: 5'- CTT' GTC' TTG' CGT' TTT' GTA' ATG-3'

30

The primers are synthesized on an DNA/RNA Synthesizer Model 394<sup>®</sup>. Thirty-five cycles are performed, each consisting of denaturing (95°C; 30 seconds), annealing (55°C; 1 minute), and extension (72°C; 90 seconds), except during the

10

15

20

25

first cycle in which the denaturing time is increased to 5 minutes, and during the last cycle in which the extension time is increased to 5 minutes.

PCR products are purified using Qia-quick<sup>®</sup> PCR purification kits (Qiagen<sup>®</sup>, cat# 28104; Chatsworth, CA). Yield and purity of the PCR product are determined spectrophotometrically at OD<sub>260</sub> on a Beckman DU 650 spectrophotometer.

# 2. <u>Dideoxy Sequence Analysis</u>

Fluorescent dye is attached to PCR products for automated sequencing using the Taq Dye Terminator Kit (Perkin-Elmer<sup>®</sup> cat# 401628). DNA sequencing is performed in both forward and reverse directions on an Applied Biosystems, Inc. (ABI) Foster City, CA., automated sequencer (Model 377). The software used for analysis of the resulting data is "Sequence Navigator<sup>®</sup>" purchased through ABI. The BRCA2<sup>(omi 1-5)</sup> sequence is entered into the Sequence Navigator software as the Standard for comparison. The Sequence Navigator software compares the sample sequence to the BRCA2<sup>(omi)</sup> standard, base by base. The Sequence Navigator highlights all differences between the BRCA2<sup>(omi)</sup> normal DNA sequence and the patient's sample sequence.

A first technologist checks the computerized results by comparing visually the BRCA2<sup>(omi 1-5)</sup> standard against the patient's sample, and again highlights any differences between the standard and the sample. The first primary technologist then interprets the sequence variations at each position along the sequence. Chromatograms from each sequence variation are generated by the Sequence Navigator and printed on a color printer. The peaks are interpreted by the first primary technologist and a second primary technologist. A secondary technologist then reviews the chromatograms. The results are finally interpreted by a geneticist. In each instance, a sequence variation is compared to known normal polymorphisms for position and base change. The ten frequent polymorphisms which occur in BRCA2 are:

- AAT (Asn) and CAT (His) at position 1093 (codon 289)
- <u>CAT</u> (His) and <u>AAT</u> (Asn) at position 1342 (codon 372)
- TCA (Ser) and TCG (Ser) at position 1593 (codon 455)

- CAT (His) and CAC (His) at position 2457 (codon 743)
- GTA (Val) and ATA (Ile) at position 2908 (codon 894)
- AAC (Asn) and GAC (Asp) at position 3199 (codon 991)
- AAA (Lys) and AAG (Lys) at position 3624 (codon 1132)
- GTT (Val) and GTC (Val) at position 4035 (codon 1269)
- TCA (Ser) and TCG (Ser) at position 7470 (codon 2414)
- GCC (Ala) and ACC (Thr) at position 9079 (codon 2951)

# 10 3. Results

5

15

20

25

30

Using the above PCR amplification and standard fluorescent sequencing technology, the 6174delT mutation may be found. Mutations are noted by the length of non-matching sequence variation. Such a lengthy mismatch pattern occurs with deletions and insertions. This mutation is named in accordance with the suggested nomenclature for naming mutations, Beaudet, A *et al.*, Human Mutation 2:245-248, (1993). The 6174delT mutation at codon 1982 of the BRCA2 gene lies in segment "Q" of exon 11. The DNA sequence results demonstrate the presence of a one base pair deletion of a T at nucleotide 6174 of the BRCA2 transcript, resulting in the appearance of an in-frame terminator (TAG) at codon position 2003. This mutation is, therefore, predicted to result in a truncated, and most likely, non-functional protein.

## **EXAMPLE 4**

# GENERATION OF MONOCLONAL AND POLYCLONAL ANTIBODIES USING GST-BRCA2 FUSION PROTEIN AS AN IMMUNOGEN

DNA primers are used to amplify a fragment of BRCA2 using PCR technology. The product is then digested with suitable restriction enzymes and fused in frame with the gene encoding glutathione S-transferase (GST) in Escherichia coli using GST expression vector pGEX (Pharmacia Biotech Inc.) The expression of the fusion protein is induced by the addition of isopropyl- $\beta$ -thiogalactopyranoside. The bacteria are then lysed and the overexpressed fusion protein is purified with glutathione-sepharose beads. The fusion protein is then verified by SDS/PAGE gel and N-terminus protein sequencing. The purified protein

10

15

20

25

30

is used to immunize rabbits according to standard procedures described in Harlow & Lane (1988). Polycolonal antibody is collected from the serum several weeks after and purified using known methods in the art. Monoclonal antibodies against all or fragments of BRCA2 protein, polypeptides, or functional equivalents are obtained using hybridoma technology, see also Harlow & Lane (1988). The BRCA2 protein or polypeptide is coupled to the carrier keyhole limpet hemocyanin in the presence of glutaraldehyde. The conjugated immunogen is mixed with an adjuvant and injected into rabbits. Spleens from antibody-containing rabbits are removed. The B-cells isolated from spleen are fused to myeloma cells using polyethylene glycol (PEG) to promote fusion. The hybrids between the myeloma and B-cells are selected and screened for the production of antibodies to immunogen BRCA2 protein or polypeptide. Positive cells are recloned to generate monoclonal antibodies.

#### **EXAMPLE 5**

### **DETECTION OF BRCA2 EXPRESSION IN HUMAN TISSUES AND CELL LINES**

The expression of BRCA2 in human tissues is determined using Northern blot analysis. Human tissues include those from pancreas, testis, prostate, ovary, breast, small intestine, and colon are obtained from Clontech Laboratories, Inc., Palo Alto, CA. The poly(A)+ mRNA Northern blots from different human tissues is hybridized to BRCA2 cDNA probes according to manufacture protocol. The expression level is further conformed by RT-PCR using oligo-d(T) as a primer and other suitable primers.

For Northern Blot analysis of cancer cell lines, the human ovarian cancer cell line SKOV-3 and the human breast cancer cell line MCF-7 are obtained from the American Type Culture Collection. Total RNA is prepared by lysing cell in the presence of guanidinium isocyanate. Poly(A)<sup>+</sup> mRNA is isolated using the PolyATract mRNA isolation system from Promega, Madison, WI. The isolated RNA is then electrophoresed under denaturing conditions and transferred to Nylon membrane. The probe used for Northern blot is a fragment of BRCA2 sequence obtained by PCR amplification. The probes are labeled with [ $\alpha$ -32P] dCTP using a random-primed labeling kit (Amersham Life Science, Arlington Heights, IL).

## **EXAMPLE 6**

5

10

15

20

25

30

## **EXPRESSION OF THE BRCA2 PROTEIN**

The whole-cell extracts of BRCA2 transfected cells are subjected to immunoprecipitation and immunoblotting to determine the BRCA2 protein level. The BRCA2 protein or polypeptide is immunoprecipitated using anti-BRCA2 antibodies prepared according to Example 4. Samples are then fractionated using SDS/PAGE gel and transferred to nitrocellulose. Western blot of the BRCA2 protein or polypeptide is performed with the indicated antibodies. Antibody reaction is revealed using enhanced chemiluminescence reagents (Dupont New England Nuclear, Boston, MA).

#### **EXAMPLE 7**

# USE OF THE BRCA2(omi1-5) GENE THERAPY

The growth of ovarian or breast cancer may be arrested by increasing the expression of the BRCA2 gene where inadequate expression of that gene is responsible for hereditary ovarian or breast cancer. Gene therapy may be performed on a patient to reduce the size of a tumor. The LXSN vector may be transformed with a BRCA2<sup>(omi1-5)</sup> coding sequence as presented SEQ ID NO:4, 6, 8, 10, or 12 or a fragment thereof.

# Vector

The LXSN vector is transformed with a fragment of the wildtype BRCA2<sup>(omi1-5)</sup> coding sequence as set forth in SEQ ID NO:4, 6, 8, 10, or 12. The LXSN-BRCA2<sup>(omi1-5)</sup> retroviral expression vector is constructed by cloning a *Sal* I linkered BRCA2<sup>(omi1-5)</sup> cDNA or fragments thereof into the *Xho* I site of the vector LXSN. Constructs are confirmed by DNA sequencing. See Holt et al., *Nature Genetics* 12: 298-302 (1996). Retroviral vectors are manufactured from viral producer cells using serum free and phenol-red free conditions and tested for sterility, absence of specific pathogens, and absence of replication-competent retrovirus by standard assays. Retrovirus is stored frozen in aliquots which have been tested.

Patients receive a complete physical exam, blood, and urine tests to determine overall health. They may also have a chest X-ray, electrocardiogram, and appropriate radiologic procedures to assess tumor stage.

Patients with metastatic ovarian cancer are treated with retroviral gene therapy by infusion of recombinant LXSN-BRCA2<sup>(omi1-5)</sup> retroviral vectors into peritoneal sites containing tumor, between 10<sup>9</sup> and 10<sup>10</sup> viral particles per dose. Blood samples are drawn each day and tested for the presence of retroviral vector by sensitive polymerase chain reaction (PCR)-based assays. The fluid which is removed is analyzed to determine:

- 1. The percentage of cancer cells which are taking up the recombinant LXSN-BRCA2<sup>(omi1-5)</sup> retroviral vector combination. Successful transfer of BRCA1 gene into cancer cells has been shown by both RT-PCR analysis and *in situ* hybridization. RT-PCR is performed with by the method of Thompson et al., *Nature Genetics* 9: 444-450 (1995), using primers derived from a BRCA2<sup>(omi1-5)</sup> coding sequence as in SEQ ID NO:4, 6, 8, 10, or 12 or fragments thereof. Cell lysates are prepared and immunoblotting is performed by the method of Jensen *et al.*, *Nature Genetics* 12: 303-308 (1996) and Jensen *et al.*, *Biochemistry* 31: 10887-10892 (1992).
- 2. Presence of programmed cell death using APOTAG® in situ apoptosis detection kit (ONCOR, INC., Gaithersburg, Maryland) and DNA analysis.
- 3. Measurement of BRCA2 gene expression by slide immunofluorescence or Western blot.

Patients with measurable disease are also evaluated for a clinical response to LXSN-BRCA2<sup>(omi1-5)</sup> especially those that do not undergo a palliative intervention immediately after retroviral vector therapy. Fluid cytology, abdominal girth, CT scans of the abdomen, and local symptoms are followed.

For other sites of disease, conventional response criteria are used as follows:

- 1. Complete Response (CR), complete disappearance of all measurable lesions and of all signs and symptoms of disease for at least 4 weeks.
- 2. Partial Response (PR), decrease of at least 50% of the sum of the products of the 2 largest perpendicular diameters of all measurable lesions as determined by 2 observations not less than 4 weeks apart. To be considered a PR, no new lesions should have appeared during this period and none should have increased in size.
- 3. Stable Disease, less than 25% change in tumor volume from previous evaluations.

25

30

5

10

15

15

20

25

4. Progressive Disease, greater than 25% increase in tumor measurements from prior evaluations. The number of doses depends upon the response to treatment.

#### 5 **EXAMPLE 8**

# PROTEIN REPLACEMENT THERAPY

Therapeutically elevated level of functional BRCA2 protein may alleviate the absence or reduced endogenous BRCA2 tumor suppressing activity. Breast or ovarian cancer is treated by the administration of a therapeutically effective amount of the BRCA2 protein, a polypeptide, or its functional equivalent in a pharmaceutically acceptable carrier. Clinically effective delivery method is applied either locally at the site of the tumor or systemically to reach other metastasized locations with known protocols in the art. These protocols may employ the methods of direct injection into a tumor or diffusion using time release capsule. A therapeutically effective dosage is determined by one of skill in the art.

Breast or ovarian cancer may be prevented by the administration of a prophylactically effective amount of the BRCA2 protein, polypeptide, or its functional equivalent in a pharmaceutically acceptable carrier. Individuals with known risk for breast or ovarian cancer are subjected to protein replacement therapy to prevent tumorigenesis or to decrease the risk of cancer. Elevated risk for breast and ovarian cancer includes factors such as carriers of one or more known BRCA1 and BRCA2 mutations, late child bearing, early onset of menstrual period, late occurrence of menopause, and certain high risk dietary habits. Clinically effective delivery method is used with known protocols in the art, such as administration into peritoneal cavity, or using an implantable time release capsule. A prophylactically effective dosage is determined by one of skill in the art.

## TABLE OF REFERENCES

- 1. Sanger, F., et al., J. Mol. Biol. 42:1617, (1980).
- 2. Beaucage, et al., Tetrahedron Letters 22:1859-1862, (1981).
  - 3. Maniatis, et al. in Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, NY, p 280-281, (1982).
  - 4. Conner, et al., Proc. Natl. Acad. Sci. U.S.A. 80:278, (1983)
  - 5. Saiki, et.al.., *Bio/Technology* 3:1008-1012, (1985)

30

- 6. Landgren, et al., Science 241:1007,(1988)
- 7. Landgren, et al., Science 242:229-237, (1988).
- 8. PCR. A Practical Approach, ILR Press, Eds. M. J. McPherson, P. Quirke, and G. R. Taylor, (1992).
- 9. Easton et al., American Journal of Human Genetics 52:678-701, (1993).
- 10. Patent No. 4,458,066.
- 11. Rowell, S., et al., American Journal of Human Genetics 55:861-865, (1994)
- 12. Miki, Y. et al., Science 266:66-71, (1994).
- 10 13. Wooster, R. et al., Science <u>265</u>:2088-2090, (1994).
  - 14. Wooster, R. et al., Nature 378:789-792, (1995).
  - 15. Beaudet, A et al., Human Mutation 2:245-248, (1993).
  - 16. Friend, S. et al. Nature Genetics 11:238, (1995).
  - 17. Teng et al, Nature Genetics 13: 241-244 (1996).
- 15 18. Couch et al, Nature Genetics <u>13</u>: 123-125 (1996).
  - 19. Tartigan et al, Nature Genetics 12: 333-337 (1996).
  - 20. Phelan et al, Nature Genetics 13: 120-122 (1996).
  - 21. Schubert et al, American Journal of Human Genetics 60: 1031-1040 (1996).
  - 22. Sambrook, et al., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1989).
  - 23. Bertwistle and Ashworth, Curr. Opin. Genet. Dev. 8(1): 14-20 (1998).
  - 24. Zhang et al., Cell 92:433-436 (1998).
  - 25. Sharan et al., Nature 386:804-810 (1997).
  - 26. Katagiri et al., Genes, Chromosomes & Cancer 21:217-222 (1988).
- 25 27. Crooke, Annu. Rev. Pharmacol. Toxicol. 32:329-376 (1992)
  - 28. Robinson-Benion and Holt, Methods Enzymol. 254:363-375 (1995).
  - 29. Harlow & Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1988.
  - 30. Shuldiner, et al., Handbook of Techniques in Endocrine Research, p. 457-486, DePablo, F., Scanes, C., eds., Academic Press, Inc., 1993.
  - 31. Holt et al., Nature Genetics 12: 298-302 (1996).
  - 32. Thompson et al., Nature Genetics 9: 444-450 (1995).
  - 33. Jensen et al., Nature Genetics <u>12</u>: 303-308 (1996)
  - 34. Jensen et al., Biochemistry 31: 10887-10892 (1992).

Although the invention has been described with reference to the presently preferred embodiments, it should be understood that various modifications can be made without departing from the spirit of the invention. Accordingly, the invention is limited only by the following claims.

(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 50 base pairs

5

60

SEQUENCE LISTING

5	(B) TYPE: nucleic acid (C) STRANDEDNESS: double (D) TOPOLOGY: linear	
3	<pre>(ii) MOLECULE TYPE: Genomic DNA (ix) FEATURE:</pre>	
10	<ul><li>(A) NAME/KEY: exon</li><li>(B) LOCATION: 150</li><li>(D) OTHER INFORMATION: Exon 5</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:	
15	TCCTGTTGTT CTACAATGTA CACATGTAAC ACCACAAAGA GATAAGTCAG	50
	(2) INFORMATION FOR SEQ ID NO:2:	
20	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 182 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: double</li><li>(D) TOPOLOGY: linear</li></ul>	
25	<pre>(ii) MOLECULE TYPE: Genomic DNA (ix) FEATURE:</pre>	
30	<ul><li>(A) NAME/KEY: exon</li><li>(B) LOCATION: 1182</li><li>(D) OTHER INFORMATION: Exon 15</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:	
35	ATTTAATTAC AAGTCTTCAG AATGCCAGAG ATATACAGGA TATGCGAATT AAGAAGAAAC AAAGGCAACG CGTCTTTCCA CAGCCAGGCA GTCTGTATCT TGCAAAAACA TCCACTCTGC CTCGAATCTC TCTGAAAGCA GCAGTAGGAG GCCAAGTTCC CTCTGCGTGT TCTCATAAAC AG	60 120 180 182
40	(2) INFORMATION FOR SEQ ID NO:3:	
	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 188 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: double</li></ul>	
45	(D) TOPOLOGY: linear	
	(ii) MOLECULE TYPE: Genomic DNA (ix) FEATURE:	
50	<ul><li>(A) NAME/KEY: exon</li><li>(B) LOCATION: 1188</li><li>(D) OTHER INFORMATION: Exon 16</li></ul>	
55	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:	
	CTGTATACGT ATGGCGTTTC TAAACATTGC ATAAAAATTA ACAGCAAAAA TGCAGAGTCT TTTCAGTTTC ACACTGAAGA TTATTTTGGT AAGGAAAGTT TATGGACTGG AAAAGGAATA CAGTTGGCTG ATGGTGGATG GCTCATACCC TCCAATGATG GAAAGGCTGG AAAAGAAGAA TTTTATAG	60 120 180
60	(2) INFORMATION FOR SEQ ID NO:4:	188
	(a) intoldinition for bug in No.4.	

5	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 10485 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
10	<pre>(ii) MOLECULE TYPE: cDNA (ix) FEATURE:  (A) NAME/KEY: Coding Sequence (B) LOCATION: 22910482</pre>	
15	(D) OTHER INFORMATION: BRCA2 (OMI1)  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:	
20	GGTGGCGCA GCTTCTGAAA CTAGGCGGCA GAGGCGGAGC CGCTGTGGCA CTGCTGCGCC TCTGCTGCGC CTCGGGTGTC TTTTGCGGCG GTGGGTCGCC GCCGGGAGAA GCGTGAGGGG ACAGATTTGT GACCGGCGCG GTTTTTGTCA GCTTACTCCG GCCAAAAAAG AACTGCACCT CTGGAGCGGA CTTATTTACC AAGCATTGGA GGAATATCGT AGGTAAAA ATG CCT ATT  Met Pro Ile  1	60 120 180 237
25	GGA TCC AAA GAG AGG CCA ACA TTT TTT GAA ATT TTT AAG ACA CGC TGC Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys Thr Arg Cys 5 10 15	285
30	AAC AAA GCA GAT TTA GGA CCA ATA AGT CTT AAT TGG TTT GAA GAA CTT Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe Glu Glu Leu 20 25 30 35	333
35	TCT TCA GAA GCT CCA CCC TAT AAT TCT GAA CCT GCA GAA GAA TCT GAA Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu Glu Ser Glu 40 45 50	381
	CAT AAA AAC AAC AAT TAC GAA CCA AAC CTA TTT AAA ACT CCA CAA AGG His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr Pro Gln Arg 55 60 65	429
40	AAA CCA TCT TAT AAT CAG CTG GCT TCA ACT CCA ATA ATA TTC AAA GAG Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile Phe Lys Glu 70 75 80	477
45	CAA GGG CTG ACT CTG CCG CTG TAC CAA TCT CCT GTA AAA GAA TTA GAT Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys Glu Leu Asp 85 90 95	525
50	AAA TTC AAA TTA GAC TTA GGA AGG AAT GTT CCC AAT AGT AGA CAT AAA Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser Arg His Lys 100 115	573
	AGT CTT CGC ACA GTG AAA ACT AAA ATG GAT CAA GCA GAT GAT GTT TCC Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp Asp Val Ser 120 125 130	621
55	TGT CCA CTT CTA AAT TCT TGT CTT AGT GAA AGT CCT GTT GTT CTA CAA Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val Val Leu Gln 135 140 145	669
60	TGT ACA CAT GTA ACA CCA CAA AGA GAT AAG TCA GTG GTA TGT GGG AGT Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val Cys Gly Ser 150 155 160	717

5			CAT His														765
10			GAA Glu														813
			TTA Leu														861
15	AGA Arg	AAT Asn	GAA Glu	GAA Glu 215	GCA Ala	TCT Ser	GAA Glu	ACT Thr	GTA Val 220	TTT Phe	CCT Pro	CAT His	GAT Asp	ACT Thr 225	ACT Thr	GCT Ala	909
20			AAA Lys 230														957
25			TTT Phe														1005
30			GCA Ala														1053
30			AGC Ser														1101
35	GAA Glu	GAT Asp	GAA Glu	GTA Val 295	TAT Tyr	GAA Glu	ACA Thr	GTT Val	GTA Val 300	GAT Asp	ACC Thr	TCT Ser	GAA Glu	GAA Glu 305	GAT Asp	AGT Ser	1149
40			TTA Leu 310														1197
45			AGC Ser														1245
50	GAA Glu 340	TGT Cys	GAA Glu	AAA Lys	TCT Ser	AAA Lys 345	AAC Asn	CAA Gln	GTG Val	AAA Lys	GAA Glu 350	AAA Lys	TAC Tyr	TCA Ser	TTT Phe	GTA Val 355	1293
	TCT Ser	GAA Glu	GTG Val	GAA Glu	CCA Pro 360	AAT Asn	GAT Asp	ACT Thr	GAT Asp	CCA Pro 365	TTA Leu	GAT Asp	TCA Ser	AAT Asn	GTA Val 370	GCA Ala	1341
55			AAG Lys														1389
60			TCT Ser 390														1437

5		GGA Gly 405															1485
3		CAA Gln															1533
10		AAA Lys															1581
15		CCA Pro															1629
20		GAT Asp															1677
25		AAG Lys 485															1725
		ATC Ile															1773
30		AAT Asn															1821
35		ACT Thr															1869
40		AAG Lys															1917
45		GCC Ala 565															1965
		TCC Ser															2013
50		GAA Glu															2061
55		CTA Leu															2109
60		CTT Leu															2157
	AAA	AGA	AGC	TGT	TCA	CAG	AAT	GAT	TCT	GAA	GAA	CCA	ACT	TTG	TCC	TTA	2205

	Lys	Arg 645	Ser	Cys	Ser	Gln	Asn 650	Asp	Ser	Glu	Glu	Pro 655	Thr	Leu	Ser	Leu	
5		AGC Ser															2253
10		TCT Ser															2301
15		TGT Cys															2349
20		CTG Leu															2397
20		AAA Lys 725															2445
25		GTA Val															2493
30		AAA Lys															2541
35		CCT Pro															2589
40		AAA Lys															2637
40		TCT Ser 805															2685
45		GTA Val															2733
50		GAA Glu															2781
55		CAA Gln															2829
60		TCA Ser															2877
		GAC Asp															2925

885 890 895

5		CTT Leu														2973
10		GTA Val														3021
15		ACA Thr														3069
10		TAT Tyr														3117
20		ATG Met 965														3165
25		AAA Lys														3213
30		TTA Leu		Pro					Ser					Phe		3261
35		TCA Ser	Asn					Leu					Ile			3309
33		ATG Met					Ile					Pro				3357
40	Cys	GTT Val 1045				Asn					Asp					3405
45		AAG Lys			Ser					Ser					Ser	3453
50		GTT Val		Ser					Ser					Gln		3501
55		TCC Ser	Lys					Ser					Thr			3549
<i>J J</i>		GCA Ala					Leu					Glu				3597
60	Gln	TTT Phe 1125				Gln					Ser					3645

5			Asn Gln Met	ACT ATC TTA Thr Ile Leu 1150	3693
10		g Asp Ala A		GTC ATA ATG Val Ile Met	3741
				TTT GAA GGT Phe Glu Gly	3789
15	Ile Lys	e Ala Gly I		AAT GAC TGT Asn Asp Cys 1200	3837
20				GTG GGG TTT Val Gly Phe 1215	3885
25			Leu Asn Val	TCT ACT GAA Ser Thr Glu 1230	3933
30		ı Phe Ser A		AAT ATT AGT Asn Ile Ser	3981
30				TCA AGT AAA Ser Ser Lys	4029
35	Ser Val	Phe Lys 1		CAT AAT GAT His Asn Asp 1280	4077
40				TTA CAA AAT Leu Gln Asn 1295	4125
45			Glu Glu Ile	ACT GAA AAT Thr Glu Asn 1310	4173
50		ı Asp Asn I		GCT GCC AGT Ala Ala Ser	4221
30				AGT AAA AAT Ser Lys Asn	4269
55	Cys Ile	Glu Thr A		TTT ACT GAT Phe Thr Asp 1360	4317
60				AAG GAG GGA Lys Glu Gly 1375	4365

5				GAT Asp	Leu					Phe					Lys		4413
				TGT Cys					Ser					Leu			4461
10			Thr	GAG Glu 1415				Lys					Ser				4509
15		Gln		GCA Ala			Lys					Ala					4557
20	Asn			GTA Val		Phe					Pro						4605
25				AAT Asn	Ser					Asp					Lys		4653
				AGT Ser					Asp					Lys			4701
30			Ser	GTC Val 1495				Thr					Val				4749
35		Gln		GAA Glu			Glu					Pro					4797
40	Phe	His 1525	Thr	GCT Ala	Ser	Gly 1	Lys 1530	Lys	Val	Lys	Ile	Ala 1535	Lys	Glu	Ser	Leu	4845
45	Asp 1540	Lys	Val	AAA Lys	Asn	Leu L545	Phe	Asp	Glu	Lys	Glu L550	Gln	Gly	Thr	Ser	Glu 1555	4893
				TTT Phe					Ala					Tyr			4941
50	Ala	Cys	r T	GAC Asp 1575	Leu	Glu	Leu	Ala 1	Cys .580	Glu	Thr	Ile	Glu 2	Ile L585	Thr	Ala	4989
55		Pro		TGT Cys			Met					Asn					5037
60	Leu			ATT Ile		Thr					Lys						5085
	TTA	TGT	AGA	CAA	ACT	GAA	AAT	CTC	AAA	ACA	TCA	AAA	AGT	ATC	TTT	TTG	5133

	Leu 1620	Cys	Arg	Gln		Glu 1625	Asn	Leu	Lys		Ser 1630	Lys	Ser	Ile	Phe	Leu 1635	
5				Val					Glu					Lys	AGT Ser 1650		5181
10			Cys					Ser					Ile		AAT Asn		5229
15		Leu					Ser					Thr			AGT Ser		5277
2.0	Thr					Ala					Arg				TTT Phe		5325
20					Arg					Asp					TAT Tyr		5373
25				Asn					Ile					Lys	AAT Asn L730		5421
30			Glu					Tyr					Ser		TCT Ser		5469
35		Tyr					Asp					Asp			TAT Tyr		5517
40	Ser					Asp					Pro				AAT Asn		5565
40					Asn					Lys					GTA Val		5613
45				Ala					Val					Cys	GTT Val 1810		5661
50			Val					Pro					Asn		GCC Ala		5709
55		Leu					Ser					Val			CCT Pro		5757
60	Phe					Gly					Val				ACA Thr		5805
00															ATT Ile		5853

CAA AAT GTA TCA AAA ATA CTT CCT CGT GTT GAT AAG AGA AAC CCA GAG

Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg Asn Pro Glu

	5				Asn					Lys					Glu	TTT Phe 2130	6621
	1.0			Asn					Glu					Glu		AAT Asn	6669
	10		Ile					Tyr					Gln			AAA Lys	6717
	15	Gln					Thr					Val				CAT His	6765
	20	TTG Leu 2180				Gln					Asn						6813
	25				Thr					Pro					Ile	GAA Glu 2210	6861
	30			Thr					Ser					Glu		GAA Glu	6909
ik,	30		Glu					Phe					Glu			GAT Asp	6957
Ond South Some Source come South	35	Lys					Ala					Phe				GAA Glu	7005
in Sand	40					Leu					Ile					GGA Gly	7053
	45				Leu	Val	Gly	Glu	Pro	Ser	Ile	Lys	Arg		Leu	TTA Leu 2290	7101
	50			Asp					Asn					Leu		GCT Ala	7149
	20		Ser					Thr					Arg			ATG Met	7197
	55	His					Pro					Pro				ACT Thr	7245
	60					Ile					Phe					CAA Gln	7293

_	TTT CTG TCT AAA TCT CAT TTG TAT GAA CAT CTG ACT TTG GAA AAA TCT Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu Glu Lys Ser 2360 2365 2370	7341
5	TCA AGC AAT TTA GCA GTT TCA GGA CAT CCA TTT TAT CAA GTT TCT GCT Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln Val Ser Ala 2375 2380 2385	7389
10	ACA AGA AAT GAA AAA ATG AGA CAC TTG ATT ACT ACA GGC AGA CCA ACC Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly Arg Pro Thr 2390 2395 2400	7437
15	AAA GTC TTT GTT CCA CCT TTT AAA ACT AAA TCA CAT TTT CAC AGA GTT Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe His Arg Val 2405 2410 2415	7485
20	GAA CAG TGT GTT AGG AAT ATT AAC TTG GAG GAA AAC AGA CAA AAG CAA Glu Gln Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg Gln Lys Gln 2420 2435	7533
25	AAC ATT GAT GGA CAT GGC TCT GAT GAT AGT AAA AAT AAG ATT AAT GAC Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys Ile Asn Asp 2440 2445 2450	7581
25	AAT GAG ATT CAT CAG TTT AAC AAA AAC AAC TCC AAT CAA GCA GCT Asn Glu Ile His Gln Phe Asn Lys Asn Asn Ser Asn Gln Ala Ala 2455 2460 2465	7629
30	GTA ACT TTC ACA AAG TGT GAA GAA GAA CCT TTA GAT TTA ATT ACA AGT Val Thr Phe Thr Lys Cys Glu Glu Glu Pro Leu Asp Leu Ile Thr Ser 2470 2475 2480	7677
35	CTT CAG AAT GCC AGA GAT ATA CAG GAT ATG CGA ATT AAG AAG AAA CAA Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys Lys Gln 2485 2490 2495	7725
40	AGG CAA CGC GTC TTT CCA CAG CCA GGC AGT CTG TAT CTT GCA AAA ACA Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu Ala Lys Thr 2500 2505 2510 2515	7773
45	TCC ACT CTG CCT CGA ATC TCT CTG AAA GCA GCA GTA GGA GGC CAA GTT Ser Thr Leu Pro Arg Ile Ser Leu Lys Ala Ala Val Gly Gly Gln Val 2520 2525 2530	7821
13	CCC TCT GCG TGT TCT CAT AAA CAG CTG TAT ACG TAT GGC GTT TCT AAA Pro Ser Ala Cys Ser His Lys Gln Leu Tyr Thr Tyr Gly Val Ser Lys 2535 2540 2545	7869
50	CAT TGC ATA AAA ATT AAC AGC AAA AAT GCA GAG TCT TTT CAG TTT CAC His Cys Ile Lys Ile Asn Ser Lys Asn Ala Glu Ser Phe Gln Phe His 2550 2555 2560	7917
55	ACT GAA GAT TAT TTT GGT AAG GAA AGT TTA TGG ACT GGA AAA GGA ATA Thr Glu Asp Tyr Phe Gly Lys Glu Ser Leu Trp Thr Gly Lys Gly Ile 2565 2570 2575	7965
60	CAG TTG GCT GAT GGT GGA TGG CTC ATA CCC TCC AAT GAT GGA AAG GCT Gln Leu Ala Asp Gly Gly Trp Leu Ile Pro Ser Asn Asp Gly Lys Ala 2580 2585 2590 2595	8013
	GGA AAA GAA GAA TTT TAT AGG GCT CTG TGT GAC ACT CCA GGT GTG GAT	8061

	Gly Lys Glu Glu Phe Tyr Arg Ala Leu Cys Asp Thr Pro Gly Val Asp 2600 2605 2610	
5	CCA AAG CTT ATT TCT AGA ATT TGG GTT TAT AAT CAC TAT AGA TGG ATC Pro Lys Leu Ile Ser Arg Ile Trp Val Tyr Asn His Tyr Arg Trp Ile 2615 2620 2625	8109
10	ATA TGG AAA CTG GCA GCT ATG GAA TGT GCC TTT CCT AAG GAA TTT GCT  Ile Trp Lys Leu Ala Ala Met Glu Cys Ala Phe Pro Lys Glu Phe Ala  2630 2635 2640	8157
15	AAT AGA TGC CTA AGC CCA GAA AGG GTG CTT CTT CAA CTA AAA TAC AGA Asn Arg Cys Leu Ser Pro Glu Arg Val Leu Leu Gln Leu Lys Tyr Arg 2645 2650 2655	8205
20	TAT GAT ACG GAA ATT GAT AGA AGC AGA AGA TCG GCT ATA AAA AAG ATA Tyr Asp Thr Glu Ile Asp Arg Ser Arg Arg Ser Ala Ile Lys Lys Ile 2660 2665 2670 2675	8253
20	ATG GAA AGG GAT GAC ACA GCT GCA AAA ACA CTT GTT CTC TGT GTT TCT Met Glu Arg Asp Asp Thr Ala Ala Lys Thr Leu Val Leu Cys Val Ser 2680 2685 2690	8301
25	GAC ATA ATT TCA TTG AGC GCA AAT ATA TCT GAA ACT TCT AGC AAT AAA Asp Ile Ile Ser Leu Ser Ala Asn Ile Ser Glu Thr Ser Ser Asn Lys 2695 2700 2705	8349
30	ACT AGT AGT GCA GAT ACC CAA AAA GTG GCC ATT ATT GAA CTT ACA GAT Thr Ser Ser Ala Asp Thr Gln Lys Val Ala Ile Ile Glu Leu Thr Asp 2710 2715 2720	8397
35	GGG TGG TAT GCT GTT AAG GCC CAG TTA GAT CCT CCC CTC TTA GCT GTC Gly Trp Tyr Ala Val Lys Ala Gln Leu Asp Pro Pro Leu Leu Ala Val 2725 2730 2735	8445
40	TTA AAG AAT GGC AGA CTG ACA GTT GGT CAG AAG ATT ATT CTT CAT GGA Leu Lys Asn Gly Arg Leu Thr Val Gly Gln Lys Ile Ile Leu His Gly 2740 2745 2750 2755	8493
40	GCA GAA CTG GTG GGC TCT CCT GAT GCC TGT ACA CCT CTT GAA GCC CCA Ala Glu Leu Val Gly Ser Pro Asp Ala Cys Thr Pro Leu Glu Ala Pro 2760 2765 2770	8541
45	GAA TCT CTT ATG TTA AAG ATT TCT GCT AAC AGT ACT CGG CCT GCT CGC Glu Ser Leu Met Leu Lys Ile Ser Ala Asn Ser Thr Arg Pro Ala Arg 2775 2780 2785	8589
50	TGG TAT ACC AAA CTT GGA TTC TTT CCT GAC CCT AGA CCT TTT CCT CTG  Trp Tyr Thr Lys Leu Gly Phe Phe Pro Asp Pro Arg Pro Phe Pro Leu  2790 2795 2800	8637
55	CCC TTA TCA TCG CTT TTC AGT GAT GGA GGA AAT GTT GGT TGT GAT Pro Leu Ser Ser Leu Phe Ser Asp Gly Gly Asn Val Gly Cys Val Asp 2805 2810 2815	8685
60	GTA ATT ATT CAA AGA GCA TAC CCT ATA CAG TGG ATG GAG AAG ACA TCA Val Ile Ile Gln Arg Ala Tyr Pro Ile Gln Trp Met Glu Lys Thr Ser 2820 2825 2830 2835	8733
30	TCT GGA TTA TAC ATA TTT CGC AAT GAA AGA GAG GAA GAA AAG GAA GCA Ser Gly Leu Tyr Ile Phe Arg Asn Glu Arg Glu Glu Glu Lys Glu Ala	8781

2840 2845 2850

5	GCA AAA TAT GTG GAG GCC CAA CAA AAG AGA CTA GAA GCC TTA TTC ACT Ala Lys Tyr Val Glu Ala Gln Gln Lys Arg Leu Glu Ala Leu Phe Thr 2855 2860 2865	8829
10	AAA ATT CAG GAG GAA TTT GAA GAA CAT GAA GAA AAC ACA ACA AAA CCA Lys Ile Gln Glu Glu Phe Glu Glu His Glu Glu Asn Thr Thr Lys Pro 2870 2875 2880	8877
15	TAT TTA CCA TCA CGT GCA CTA ACA AGA CAG CAA GTT CGT GCT TTG CAA Tyr Leu Pro Ser Arg Ala Leu Thr Arg Gln Gln Val Arg Ala Leu Gln 2885 2890 2895	8925
13	GAT GGT GCA GAG CTT TAT GAA GCA GTG AAG AAT GCA GCA GAC CCA GCT Asp Gly Ala Glu Leu Tyr Glu Ala Val Lys Asn Ala Ala Asp Pro Ala 2900 2905 2910 2915	8973
20	TAC CTT GAG GGT TAT TTC AGT GAA GAG CAG TTA AGA GCC TTG AAT AAT Tyr Leu Glu Gly Tyr Phe Ser Glu Glu Gln Leu Arg Ala Leu Asn Asn 2920 2925 2930	9021
25	CAC AGG CAA ATG TTG AAT GAT AAG AAA CAA GCT CAG ATC CAG TTG GAA His Arg Gln Met Leu Asn Asp Lys Lys Gln Ala Gln Ile Gln Leu Glu 2935 2940 2945	9069
30	ATT AGG AAG GCC ATG GAA TCT GCT GAA CAA AAG GAA CAA GGT TTA TCA Ile Arg Lys Ala Met Glu Ser Ala Glu Gln Lys Glu Gln Gly Leu Ser 2950 2955 2960	9117
35	AGG GAT GTC ACA ACC GTG TGG AAG TTG CGT ATT GTA AGC TAT TCA AAA Arg Asp Val Thr Thr Val Trp Lys Leu Arg Ile Val Ser Tyr Ser Lys 2965 2970 2975	9165
33	AAA GAA AAA GAT TCA GTT ATA CTG AGT ATT TGG CGT CCA TCA TCA GAT Lys Glu Lys Asp Ser Val Ile Leu Ser Ile Trp Arg Pro Ser Ser Asp 2980 2985 2990 2995	9213
40	TTA TAT TCT CTG TTA ACA GAA GGA AAG AGA TAC AGA ATT TAT CAT CTT Leu Tyr Ser Leu Leu Thr Glu Gly Lys Arg Tyr Arg Ile Tyr His Leu 3000 3005 3010	9261
45	GCA ACT TCA AAA TCT AAA AGT AAA TCT GAA AGA GCT AAC ATA CAG TTA Ala Thr Ser Lys Ser Lys Ser Lys Ser Glu Arg Ala Asn Ile Gln Leu 3015 3020 3025	9309
50	GCA GCG ACA AAA AAA ACT CAG TAT CAA CAA CTA CCG GTT TCA GAT GAA Ala Ala Thr Lys Lys Thr Gln Tyr Gln Gln Leu Pro Val Ser Asp Glu 3030 3035 3040	9357
EE	ATT TTA TTT CAG ATT TAC CAG CCA CGG GAG CCC CTT CAC TTC AGC AAA Ile Leu Phe Gln Ile Tyr Gln Pro Arg Glu Pro Leu His Phe Ser Lys 3045 3050 3055	9405
55	TTT TTA GAT CCA GAC TTT CAG CCA TCT TGT TCT GAG GTG GAC CTA ATA  Phe Leu Asp Pro Asp Phe Gln Pro Ser Cys Ser Glu Val Asp Leu Ile  3060 3075	9453
60	GGA TTT GTC GTT TCT GTG AAA AAA ACA GGA CTT GCC CCT TTC GTC Gly Phe Val Val Ser Val Val Lys Lys Thr Gly Leu Ala Pro Phe Val 3080 3085 3090	9501

5	TAT TTG TCA GAC GAA TGT TAC AAT TTA CTG GCA ATA AAG TTT TGG ATA Tyr Leu Ser Asp Glu Cys Tyr Asn Leu Leu Ala Ile Lys Phe Trp Ile 3095 3100 3105	9549
10	GAC CTT AAT GAG GAC ATT ATT AAG CCT CAT ATG TTA ATT GCT GCA AGC Asp Leu Asn Glu Asp Ile Ile Lys Pro His Met Leu Ile Ala Ala Ser 3110 3115 3120	9597
10	AAC CTC CAG TGG CGA CCA GAA TCC AAA TCA GGC CTT CTT ACT TTA TTT Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu Thr Leu Phe 3125 3130 3135	9645
15	GCT GGA GAT TTT TCT GTG TTT TCT GCT AGT CCA AAA GAG GGC CAC TTT Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu Gly His Phe 3140 3145 3150 3155	9693
20	CAA GAG ACA TTC AAC AAA ATG AAA AAT ACT GTT GAG AAT ATT GAC ATA Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn Ile Asp Ile 3160 3165 3170	9741
25	CTT TGC AAT GAA GCA GAA AAC AAG CTT ATG CAT ATA CTG CAT GCA AAT Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu His Ala Asn 3175 3180 3185	9789
2.0	GAT CCC AAG TGG TCC ACC CCA ACT AAA GAC TGT ACT TCA GGG CCG TAC Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser Gly Pro Tyr 3190 3195 3200	9837
30	ACT GCT CAA ATC ATT CCT GGT ACA GGA AAC AAG CTT CTG ATG TCT TCT Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu Met Ser Ser 3205 3210 3215	9885
35	CCT AAT TGT GAG ATA TAT TAT CAA AGT CCT TTA TCA CTT TGT ATG GCC Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu Cys Met Ala 3220 3225 3230 3235	9933
40	AAA AGG AAG TCT GTT TCC ACA CCT GTC TCA GCC CAG ATG ACT TCA AAG Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met Thr Ser Lys 3240 3245 3250	9981
45	TCT TGT AAA GGG GAG AAA GAG ATT GAT GAC CAA AAG AAC TGC AAA AAG Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn Cys Lys Lys 3255 3260 3265	10029
50	AGA AGA GCC TTG GAT TTC TTG AGT AGA CTG CCT TTA CCT CCA CCT GTT Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro Pro Pro Val 3270 3275 3280	10077
50	AGT CCC ATT TGT ACA TTT GTT TCT CCG GCT GCA CAG AAG GCA TTT CAG Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys Ala Phe Gln 3285 3290 3295	10125
55	CCA CCA AGG AGT TGT GGC ACC AAA TAC GAA ACA CCC ATA AAG AAA AAA Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile Lys Lys Lys 3300 3305 3310 3315	10173
60	GAA CTG AAT TCT CCT CAG ATG ACT CCA TTT AAA AAA TTC AAT GAA ATT Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe Asn Glu Ile 3320 3330	10221

5	Sei	Lei	ı Tro	GAA Glu 3335	Ser	AAT Asn	' TCA Ser	A ATA	A GCT Ala 3340	Asp	GAA Glu	A GAA 1 Glu	Leu	GCA Ala 3345	ı Leı	3 ATA 1 Ile	10269
J	AAT Asn	T ACC	C CAA Glr 3350	ı Ala	CTI Leu	TTG Leu	TCI Ser	GGT Gly 3355	/ Ser	ACA Thr	GGA Gly	A GAA ⁄ Glu	AAA Lys 3360	CAA Glr	A TTI	T ATA	10317
10	TCT Ser	GTC Val	. Ser	GAA Glu	TCC Ser	Thr	AGG Arg 3370	Thr	GCT Ala	CCC Pro	ACC Thr	AGT Ser 3375	Ser	GAA Glu	A GAT L Asp	TAT Tyr	10365
15	CTC Leu 3380	Arg	CTG Leu	AAA Lys	Arg	CGT Arg 3385	TGT Cys	' ACT Thr	ACA Thr	TCT	CTG Leu 3390	Ile	AAA Lys	GAA Glu	CAG Gln	GAG Glu 3395	10413
20	AGT Ser	TCC Ser	CAG Gln	Ala	AGT Ser 3400	Thr	GAA Glu	GAA Glu	Cys	GAG Glu 3405	Lys	AAT Asn	AAG Lys	CAG Gln	GAC Asp 3410	ACA Thr	10461
25			Thr			TAT Tyr											10485
		,				ATIOI				NO:	5:						
30		(	(A) (B) (C)	LENO TYP: STR.	GTH: E: a: ANDE	CHARA 3418 mino DNESS Y: 1:	8 am aci	ino d ingl	acid	S							
35						TYPI TYPE:											
		(:	xi)	SEQUI	ENCE	DESC	CRIP	TION	: SE	Q ID	NO:	5:					
40	1				5					10		Phe			15	-	
				20					25			Ser		30			
45			35					40				Ser	45				
		50					55					Asn 60					
<b>5</b> 0	65					70					75	Ser				80	
50					85					90		Gln			95	_	
				100					105			Asn		110			
55			115					120				Met	125				
		130					135					Ser 140					
	145					150					155	Asp				160	
60	Cys	Gly	Ser	Leu	Phe 165	His	Thr	Pro	Lys	Phe 170	Val	Lys	Gly	Arg	Gln 175	Thr	
	Pro	Lvs	His	Tle	Ser	Glu	Ser	T.=11	G117	ת ו ת	C1	7/27	7) ~~~	Dwa	70	34 - L	

				180					185					190		
	Ser	Trp	Ser 195	Ser	Ser	Leu	Ala	Thr 200	Pro	Pro	Thr	Leu	Ser 205	Ser	Thr	Val
5	Leu	Ile 210	Val	Arg	Asn	Glu	Glu 215	Ala	Ser	Glu	Thr	Val 220	Phe	Pro	His	Asp
	Thr 225	Thr	Ala	Asn	Val	Lys 230	Ser	Tyr	Phe	Ser	Asn 235	His	Asp	Glu	Ser	Leu 240
10	Lys	Lys	Asn	Asp	Arg 245	Phe	Ile	Ala	Ser	Val 250	Thr	Asp	Ser	Glu	Asn 255	Thr
	Asn	Gln	Arg	Glu 260	Ala	Ala	Ser	His	Gly 265	Phe	Gly	Lys	Thr	Ser 270	Gly	Asn
	Ser	Phe	Lys 275	Val	Asn	Ser	Cys	Lys 280	Asp	His	Ile	Gly	Lys 285	Ser	Met	Pro
15	Asn	Val 290	Leu	Glu	Asp	Glu	Val 295	Tyr	Glu	Thr	Val	Val 300	Asp	Thr	Ser	Glu
	Glu 305	Asp	Ser	Phe	Ser	Leu 310	Cys	Phe	Ser	Lys	Cys 315	Arg	Thr	Lys	Asn	Leu 320
20	Gln	Lys	Val	Arg	Thr 325	Ser	Lys	Thr	Arg	330	Lys	Ile	Phe	His	Glu 335	Ala
	Asn	Ala	Asp	Glu 340	Cys	Glu	Lys	Ser	Lys 345	Asn	Gln	Val	Lys	Glu 350	Lys	Tyr
	Ser	Phe	Val 355	Ser	Glu	Val	Glu	Pro 360	Asn	Asp	Thr	Asp	Pro 365	Leu	Asp	Ser
25	Asn	Val 370	Ala	His	Gln	Lys	Pro 375	Phe	Glu	Ser	Gly	Ser 380	Asp	Lys	Ile	Ser
	385				Pro	390			_		395					400
30		=			Gly 405					410					415	
			_	420	Gln				425					430		
o =			435		Lys			440					445			
35		450			Pro	_	455		_			460				
	465		_	_	Asp	470					475					480
40					Lys 485					490					495	
				500	Ile Asn				505					510		
<i>1</i> E			515					520					525			
45		530	_		Thr		535					540				
	545				Lys Ala	550					555					560
50	_		_		565 Ser					570					575	
		_		580	Glu				585	_			_	590		_
55			595	_	Leu			600	_	_	_	_	605		_	_
		610			Leu		615	-				620				
	625				Arg	630					635					640
60				-	645 Ser		-			650	_				655	
	ъсu	Der	ыси	660	DCI	Der	1116	GIY	665	116	ыcu	arg	nys	670	DCT	2.3

	Asn	Glu	Thr 675	Cys	Ser	Asn	Asn	Thr 680	Val	Ile	Ser	Gln	Asp 685	Leu	Asp	Tyr
5	Lys	Glu 690		Lys	Cys	Asn	Lys 695		Lys	Leu	Gln	Leu 700		Ile	Thr	Pro
_	Glu 705	Ala	Asp	Ser	Leu	Ser 710	Cys	Leu	Gln	Glu	Gly 715	Gln	Cys	Glu	Asn	Asp 720
	Pro	Lys	Ser	Lys	Lys 725	Val	Ser	Asp	Ile	Lys 730	Glu	Glu	Val	Leu	Ala 735	Ala
10	Ala	Cys	His	Pro 740	Val	Gln	His	Ser	Lys 745	Val	Glu	Tyr	Ser	Asp 750	Thr	Asp
		Gln	755		_			760	_	_			765			
15		11e 770					775	_	_			780				
	785	Ser		_	_	790		_	_		795	_	_			800
3.0		Asn	_		805	_				810	_				815	
20		Asn Leu		820					825					830		
		Gln	835			-	-	840	_				845		_	_
25		850 Glu					855					860		_		
	865	Leu				870		_			875			_		880
30		Arg			885					890					895	
	Asp	Leu	Thr	900 Cys	Val	Asn	Glu	Pro	905 Ile	Phe	Lys	Asn	Ser	910 Thr	Met	Val
	Leu	Tyr	915 Gly	Asp	Thr	Gly	_	920 Lys	Gln	Ala	Thr		925 Val	Ser	Ile	Lys
35	_	930 Asp	Leu	Val	Tyr		935 Leu	Ala	Glu	Glu		940 Lys	Asn	Ser	Val	_
	945 Gln	His	Ile	Lys	Met 965	950 Thr	Leu	Gly	Gln	Asp	955 Leu	Lys	Ser	Asp	Ile 975	960 Ser
40	Leu	Asn	Ile	Asp 980		Ile	Pro	Glu	Lys 985		Asn	Asp	Tyr	Met 990		Lys
	Trp	Ala	Gly 995	Leu	Leu	Gly	Pro	Ile 1000		Asn	His	Ser	Phe 1009		Gly	Ser
45		Arg 1010	)				1019	5				1020	)			
	102			_		1030	)	_	_		103	5		_		104
50		Leu		-	1045	5				1050	)			_	1055	5
50		Lys Ser		1060	)				1069	5				1070	)	
		Met	1075	5				1080	) _	_			1089	5		
55		1090 Ser	)				1099	5				1100	)			
	110					1110	)				1119	5				112
60		Gln			1125	5				1130	)	_			1135	5
	Lys	Thr	Thr	1140 Ser		Glu	Cys	Arg	1145 Asp		Asp	Leu	His	1150 Val		Met

		1155			1160					1165			
	Asn Ala 1170	Pro Ser	Ile Gly		Val 2		Ser		Lys 1180		Phe	Glu	Gly
5	Thr Val	Glu Ile	Lys Arg 119		Phe .	Ala		Leu 1195		Lys	Asn	Asp	Cys 120
		Ser Ala	1205				1210					1215	;
10		Phe Tyr 1220	)			1225					1230	)	
	Ala Leu	Gln Lys 1235	Ala Val	_	Leu 1240		Ser	Asp	Ile	Glu 1245		Ile	Ser
	1250			1255					1260	)			
15	1265	Asp Ser	127	0				1275					128
	-	Val Ser	1285				1290	)				1295	5
20		Glu Met 130	0			1305					1310	)	
		Arg Asn 1315			1320					1325	5		
	1330			1335	i				1340	)			
25	1345	Val Cys	135	0				1355	;				136
		Asn Ile	1365				1370	)				1375	5
30		Gln Ile 138	0			1385	•				1390	)	
	-	Ala Gln 1395			1400					1405	5		
	141		-	1415	5				1420	0			
35	1425	Phe Phe	143	0				1435	5				144
		Phe Asn	1445				1450	)				145	5
40		Asn Phe 146	0			1465	5				1470	)	
	_	Met Asp			1480	ı				1489	5		
4.5	149			1495	5				150	0			
45	1505	Gln Gly	151	. 0				1515	5				152
		Gly Phe	1525				1530	)				153	5
50		Leu Asp	0			1545	5				155	0	
		Glu Ile 1555 Glu Ala			1560	)				156	5		
55	157			1575	5				158	0			
55	1585	Asn Leu	159	90				159	5				160
		Asn Leu	1605				1610	0				161	5
60		162 Leu Lys	0			1625	5				163	0	
	TIE PHE	1635	· vат шув	, vai	1640		A311	vai	Ciu	164		****	

	Lys	Ser 1650	Pro	Ala	Thr	Cys	Tyr 1655		Asn	Gln	Ser	Pro 1660	_	Ser	Val	Ile
	Glu		Ser	Δla	Leu	Ala			Thr	Ser	Cvs	Ser	Ara	Lvs	Thr	Ser
5	1665		DCI	7114	204	1670		- 1 -			1675			272		168
5			<b>a</b> 1	m³	<b>a</b>			<b>a</b> 1	n 7	T			T	7	a1	
			Gln		1685	5				1690	)				1695	5
	Ile	Phe	Asp	Gly 1700		Pro	Glu	Arg	Ile 1705		Thr	Ala	Asp	Tyr 1710		Gly
10	Asn	Tyr	Leu 1715		Glu	Asn	Asn	Ser 1720		Ser	Thr	Ile	Ala 1725		Asn	Asp
	Lys	Asn 1730	His		Ser	Glu	Lys 1735	Gln		Thr	Tyr	Leu 1740		Asn	Ser	Ser
15	Met 1745	Ser	Asn	Ser	Tyr	Ser 1750	Tyr		Ser	Asp	Glu 1755	Val		Asn	Asp	Ser 176
			Leu	Ser	Lys 1765	Asn		Leu	Asp	Ser 1770	Gly		Glu	Pro	Val 1775	Leu
	Lare	Λen	Val	Glu			Laze	Aen	Thr			Ser	Laze	Val		
2.0	-			1780	) _		-		1785	5			_	1790	)	
20			Lys 1795	5				1800	)				1805	5		
	Суѕ	Val 1810	Glu )	Glu	Leu	Val	Thr 1815		Ser	Ser	Pro	Cys 1820	-	Asn	Lys	Asn
	Ala	Ala	Ile	Lys	Leu	Ser	Ile	Ser	Asn	Ser	Asn	Asn	Phe	Glu	Val	Gly
25	1825					1830					1835					184
	Pro	Pro	Ala	Phe	Arg 1845		Ala	Ser	Gly	Lys 1850		Val	Cys	Val	Ser 1855	
	Glu	Thr	Ile	Lys 1860	Lys		Lys	Asp	Ile 1865	Phe		Asp	Ser	Phe 1870	Ser	
30	Val	Ile	Lys	Glu		Asn	Glu		Lys		Lys	Ile	_	Gln		Lys
	~ 1 _	3.4 ±	1875		<b>a</b>	m	<b>a</b> 1	1880		3	7		1885		<b>-</b> 1-	T
	TIE		Ala	GIA	Cys	Tyr			ьeu	Asp	Asp			Asp	TTE	ьeu
		1890					1895			_		1900				
			Ser	Leu	Asp			Glu	Cys	Ser			Ser	His	Lys	
35	1905					1910					1915					192
			Asp		1925	5				1930	)				1935	5
	Ser	Gly	Leu	Glu 1940	_	Val	Ser	Lys	Ile 1945		Pro	Cys	Asp	Val 1950		Leu
40	Glu	Thr	Ser 1955	_	Ile	Cys	Lys	Cys 1960		Ile	Gly	Lys	Leu 1965		Lys	Ser
	Val	Ser 1970	Ser	Ala	Asn	Thr	Cys 1975	_	Ile	Phe	Ser	Thr 1980		Ser	Gly	Lys
45	Ser 1985	Val	Gln	Val	Ser	Asp	Ala		Leu	Gln	Asn 1995	Ala		Gln	Val	Phe 200
13			Ile	Glu	Agn			Larg	Gln	Val			Lve	Val	T.e.11	
	JCI	Oiu	110	Oru	2005		1111	БУБ	OIII	2010		JCI	цуз	vai	2015	
	Lys	Ser	Asn	Glu 2020	His		Asp	Gln		Thr		Glu	Glu		Thr	
50	Ile	Arg	Thr	Pro		His	Leu	Ile 2040			Lys	Gly				Asn
	Val		2035 Asn		Ser	Ala		Ser		Phe	Ser				Gly	Lys
			Ser	Ile	Leu				Leu	His	_			Gly	Val	
55	2069 Glu		Phe	Asp	Leu	2070 Ile		Thr	Glu	His	2075 Ser		His	Tyr	Ser	208 Pro
					2085	5				2090	)				2095	5
	Thr	Ser	Arg	Gln 2100		Val	Ser	Lys	Ile 2105		Pro	Arg	Val	Asp 2110		Arg
60	Asn	Pro	Glu 2115		Cys	Val	Asn	Ser 2120	Glu		Glu	Lys	Thr 2129		Ser	Lys
	Glu	Phe	Lys		Ser	Asn	Asn			Val	Glu	Gly			Ser	Glu

		2130	)				2135	i				2140				
	Asn 2145	Asn	His	Ser	Ile	Lys 2150		Ser	Pro	Tyr	Leu 2155		Gln	Phe	Gln	Gln 216
5	-	Lys			2165	;		_		2170	l				2175	5
		His		2180	)				2185	5				2190	)	
10	Glu	Ile	Gly 2195		Thr	Glu	Thr	Phe 2200		Asp	Val	Pro	Val 2205		Thr	Asn
		Glu 2210	)	_			2215	,	_	_		2220	)			
	2225	_				2230	)	_			2235	5	_	_		224
15		Asp			2245	5				2250	1				2255	5
		Glu		2260	)				2265	5		_		2270	)	_
20	_	Gly	2275	5				2280	)				2285	5		
		Leu 2290	)			_	2295	5				2300	)			
0.5	230	_				2310	)				2315	5				232
25		Met			2325	5				2330	)				2335	5
		Thr	-	2340	)				2345	5				2350	)	
30	_	Gln	2355	5			_	2360	)		_		2365	5		
		Lys 2370	)				2375	5				2380	)			
	val 238	Ser 5	Ala	Thr	Arg	Asn 2390		ьуs	Met	_	H15		шė	Thr	Thr	G1y 240
35	Arg	Pro	Thr	Lys	Val 2405		Val	Pro	Pro	Phe 2410	_	Thr	Lys	Ser	His 2415	
		Arg		2420	)	_		_	2425	5				2430	)	
40		Lys	2435	5				2440	)				2445	5		
	Ile	Asn 2450	_	Asn	Glu	Ile	His 2455		Phe	Asn	Lys	Asn 2460		Ser	Asn	Gln
	Ala 246		Ala	Val	Thr	Phe 2470		Lys	Cys	Glu	Glu 2475		Pro	Leu	Asp	Leu 248
45	Ile	Thr	Ser	Leu	Gln 2485		Ala	Arg	Asp	Ile 2490		Asp	Met	Arg	Ile 2495	_
	Lys	Lys	Gln	Arg 2500	Gln		Val	Phe	Pro 2505	Gln		Gly	Ser	Leu 2510	_	Leu
50	Ala	Lys	Thr 2515		Thr	Leu	Pro	Arg 2520		Ser	Leu	Lys	Ala 2525		Val	Gly
	Gly	Gln 2530		Pro	Ser	Ala	Cys 2535		His	Lys	Gln	Leu 2540	_	Thr	Tyr	Gly
	254					2550	)				2555	5				256
55		Phe			2565	5				2570	)				2575	5
		Gly		2580	)				2585	5				2590	)	
60	_	Lys	2595	5	_			2600	)	_			2605	5		
	Gly	Val 261	_	Pro	Lys	Leu	Ile 2615		Arg	Ile	Trp	Val 2620	_	Asn	His	Tyr

	Arg Tr 2625	p Ile	e Ile		Lys 2630		Ala	Ala	Met	Glu 2635		Ala	Phe	Pro	Lys 264
5	Glu Ph	ne Ala	a Asn	Arg 2645		Leu	Ser	Pro	Glu 2650		Val	Leu	Leu	Gln 2655	
	Lys Ty	r Arg	Tyr 266	_	Thr	Glu	Ile	Asp 2665		Ser	Arg	Arg	Ser 2670		Ile
	Lys Ly	s Ile 26		Glu	Arg	_	Asp 2680		Ala	Ala	Lys	Thr 2685		Val	Leu
10		90	_			2695	5				2700	)			
	Ser As 2705				2710	)				2715	5				272
15	Leu Th			2725	;				2730	)				2735	5
	Leu Al		274	0				2745	5				2750	)	
	Leu Hi	27	55				2760	)				2765	5		
20	Glu Al	la Pro 770	Glu	Ser	Leu	Met 2775		Lys	Ile	Ser	Ala 2780		Ser	Thr	Arg
	Pro Al 2785		_	_	2790	)				2795	5				280
25	Phe Pr			2805	5				2810	)				2815	5
	Cys Va		282	0				2825	5				2830	)	
	Lys Th	28	35	_			2840	)				2845	5		
30		350				2855	5				2860	)			
	Leu Ph 2865				2870	)				2875	5				288
35	Thr Ly			2885	5				2890	)				2899	5
	Ala Le		290	0				2905	5				2910	)	
	Asp Pi	29	15				2920	C				292	5		
40		930				2935	5				2940	0			
	Gln Le 2945				295	0				295	5				296
45	Gly Le		_	296	5				297	C				297	5
	Tyr Se	_	298	0	_	_		298	5				2990	)	
T 0	Ser Se	29	95	_			3000	0				300	5		
50		010				3019	5				302	0			
	Ile G				303	0				303	5				304
55	Ser A			304	5				305	0				305	5
	Phe Se	_	306	0				306	5				307	)	
60	Asp Le	30	75				308	0				308	5		
60		090				309	5				310	0			
	Phe T	гъ тт	e asp	ьeu	ASI	GIU	ASP	тте	тте	цуs	PIO	птв	met	ьeu	TIE

	3105	3110	l	3115		312	
	Ala Ala Se	r Asn Leu Gln	Trp Arg Pro	Glu Ser I 3130	Lys Ser Gl	ly Leu Leu 3135	
5	Thr Leu Pho	3125 e Ala Gly Asp 3140	Phe Ser Val	Phe Ser A			
	Gly His Pho	e Gln Glu Thr					
10		e Leu Cys Asn				is Ile Leu	
	His Ala Asi 3185	n Asp Pro Lys 3190	=	Pro Thr I	Lys Asp Cy	s Thr Ser 320	
	Gly Pro Ty:	r Thr Ala Gln 3205	Ile Ile Pro	Gly Thr G	Gly Asn Ly	ys Leu Leu 3215	
15	Met Ser Se	r Pro Asn Cys 3220	Glu Ile Tyr 322			eu Ser Leu 230	
	32:		3240		3245		
20	3250	s Ser Cys Lys	3255	:	3260		
	3265	s Arg Arg Ala 3270	)	3275		328	
0.5		l Ser Pro Ile 3285	-	3290		3295	
25		n Pro Pro Arg 3300	330	5	33	310	
	33		3320		3325		
30	3330	e Ser Leu Leu	3335	:	3340		
	3345	e Asn Thr Gln 3350 e Ser Val Ser	)	3355		336	
35		3365 r Leu Arg Leu		3370		3375	
		3380 u Ser Ser Gln	338	5	3.3	390	
	33	95 r Ile Thr Thr	3400		3405	- -	
40	3410		3415				
		2) INFORMATION		NO:6:			
45	(A	SEQUENCE CHARA LENGTH: 1048 TYPE: nuclei	35 base pair	s			
	(C	) TYPE: NUCLEI ) STRANDEDNESS ) TOPOLOGY: li	S: double				
50		MOLECULE TYPE					
~ -		FEATURE:	<del>-</del>				
		A) NAME/KEY: 0 B) LOCATION: 2		nce			
55	·	D) OTHER INFOR					
		SEQUENCE DESC					
60	TCTGCTGCGC	GCTTCTGAAA CT	TTTGCGGCG GT	GGGTCGCC	GCCGGGAGA	A GCGTGAGGGG	60 120
		GACCGGCGCG GT					180 237

Met	Pro	Ile
1		

5		GAG Glu								:	285
10		GAT Asp								:	333
15		GCT Ala								,	381
20		AAC Asn 55						_			429
20		TAT Tyr									477
25		ACT Thr									525
30		TTA Leu									573
35		ACA Thr									621
40		CTA Leu 135							CAA Gln		669
40		GTA Val									717
45		ACA Thr									765
50		AGT Ser							TCA Ser 195		813
55		GCT Ala									861
60		GAA Glu 215									909
00									AAT Asn		957

230 235 240

5		ATC Ile							1005
10		AGT Ser							1053
15		TGC Cys							1101
13		GTA Val 295							1149
20		TGT Cys							1197
25		AAG Lys							1245
30		AAA Lys							1293
2.5		GAA Glu							1341
35		CCC Pro 375							1389
40		TTG Leu							1437
45		CAG Gln							1485
50		ATT Ile							1533
r r		TTT Phe							1581
55		TCA Ser 455							1629
60		GAG Glu							1677

5		GCA Ala							1725
10		AAG Lys							1773
		AGT Ser							1821
15		GCC Ala 535							1869
20		GAC Asp							1917
25		ACC Thr							1965
30		TTG Leu							2013
30		TCT Ser							2061
35		AAC Asn 615							2109
40		TTT Phe							2157
45		TGT Cys							2205
50		TTT Phe							2253
30		AAT Asn							2301
55		AAG Lys 695							2349
60		TGC Cys							2397

5		AAA Lys 725															2445
J		GTA Val															2493
10		AAA Lys															2541
15		CCT Pro															2589
20		AAA Lys															2637
25		TCT Ser 805															2685
		GTA Val															2733
30		GAA Glu															2781
35		CAA Gln															2829
40		TCA Ser															2877
45		GAC Asp 885															2925
		CTT Leu															2973
50		GTA Val															3021
55		ACA Thr															3069
60		TAT Tyr															3117
	AAA	ATG	ACT	CTA	GGT	CAA	GAT	TTA	AAA	TCG	GAC	ATC	TCC	TTG	AAT	ATA	3165

	Lys I	Met 965	Thr	Leu	Gly	Gln	Asp 970	Leu	Lys	Ser	Asp	Ile 975	Ser	Leu	Asn	Ile	
5	GAT A Asp 1 980																3213
10	CTC 1			Pro					Ser					Phe			3261
15	GCT S		Asn					Leu					Ile				3309
20	AAA A	Met					Ile					Pro					3357
20	TGT ( Cys \					Asn					Asp						3405
25	AGC A Ser I				Ser					Ser					Ser		3453
30	GTA (			Ser					Ser					Gln			3501
35	TTT :		Lys					Ser					Thr				3549
40	AAG (	Ala					Leu					Glu					3597
40	CAG : Gln :					Gln					Ser						3645
45	AGT A Ser 1				Val					Met					Thr		3693
50	TCT (			Cys					Leu					Asn			3741
55	TCG A		Gly					Ser					Gly				3789
60	ATT A	Lys					Gly					Asp					3837
00	GCT T																3885

1205 1210 1215

5	TAT TCT GCT CAT GGC ACA AAA CTG AAT GTT TCT ACT GAA GCT CTG CAA Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu Ala Leu Gln 1220 1225 1230 1235	3933
10	AAA GCT GTG AAA CTG TTT AGT GAT ATT GAG AAT ATT AGT GAG GAA ACT Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser Glu Glu Thr 1240 1245 1250	3981
15	TCT GCA GAG GTA CAT CCA ATA AGT TTA TCT TCA AGT AAA TGT CAT GAT Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Lys Cys His Asp 1255 1260 1265	4029
	TCT GTC GTT TCA ATG TTT AAG ATA GAA AAT CAT AAT GAT AAA ACT GTA Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp Lys Thr Val 1270 1275 1280	4077
20	AGT GAA AAA AAT AAT AAA TGC CAA CTG ATA TTA CAA AAT AAT ATT GAA Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn Asn Ile Glu 1285 1290 1295	4125
25	ATG ACT ACT GGC ACT TTT GTT GAA GAA ATT ACT GAA AAT TAC AAG AGA Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn Tyr Lys Arg 1300 1305 1310 1315	4173
30	AAT ACT GAA AAT GAA GAT AAC AAA TAT ACT GCT GCC AGT AGA AAT TCT Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser Arg Asn Ser 1320 1325 1330	4221
35	CAT AAC TTA GAA TTT GAT GGC AGT GAT TCA AGT AAA AAT GAT ACT GTT His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn Asp Thr Val 1335 1340 1345	4269
33	TGT ATT CAT AAA GAT GAA ACG GAC TTG CTA TTT ACT GAT CAG CAC AAC Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp Gln His Asn 1350 1355 1360	4317
40	ATA TGT CTT AAA TTA TCT GGC CAG TTT ATG AAG GAG GGA AAC ACT CAG Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly Asn Thr Gln 1365 1370 1375	4365
45	ATT AAA GAA GAT TTG TCA GAT TTA ACT TTT TTG GAA GTT GCG AAA GCT Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val Ala Lys Ala 1380 1385 1390 1395	4413
50	CAA GAA GCA TGT CAT GGT AAT ACT TCA AAT AAA GAA CAG TTA ACT GCT Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln Leu Thr Ala 1400 1405 1410	4461
55	ACT AAA ACG GAG CAA AAT ATA AAA GAT TTT GAG ACT TCT GAT ACA TTT Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser Asp Thr Phe 1415 1420 1425	4509
23	TTT CAG ACT GCA AGT GGG AAA AAT ATT AGT GTC GCC AAA GAG TCA TTT Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val Ala Lys Glu Ser Phe 1430 1435 1440	4557
60	AAT AAA ATT GTA AAT TTC TTT GAT CAG AAA CCA GAA GAA TTG CAT AAC Asn Lys Ile Val Asn Phe Phe Asp Gln Lys Pro Glu Glu Leu His Asn 1445 1450 1455	4605

5					Ser Asp		AAG AAC AAA Lys Asn Lys 1	
10		Leu Ser					CAC AAA ATA His Lys Ile 1490	
10				Gly Thr			GTG ACC TTC Val Thr Phe 1505	
15	Gly Gln					Glu Pro	ACT CTG TTG Thr Leu Leu 1520	
20			Ser Gly				AAG GAA TCT Lys Glu Ser	
25					Glu Lys		GGT ACT AGT Gly Thr Ser	
2.0		Ser Phe					AAG TAC AGA Lys Tyr Arg 1570	
30				Leu Ala			GAG ATC ACA Glu Ile Thr 1585	
35	Ala Pro					Leu Asn	AAT GAT AAA Asn Asp Lys 1600	
40		Ser Ile	Glu Thr				TTA AGT GAT Leu Ser Asp	
45					Lys Thr		AGT ATC TTT Ser Ile Phe	
50		Lys Val					GCA AAA AGT Ala Lys Ser 1650	
30				Gln Ser			ATT GAA AAT Ile Glu Asn 1665	
55	Ala Leu					Lys Thr	TCT GTG AGT Ser Val Ser 1680	
60		Leu Leu	Glu Ala				GGA ATA TTT Gly Ile Phe	

5	GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT TTG Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr Leu 1700 1705 1710 1715	5373
J	TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT CAT Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn His 1720 1725 1730	5421
10	CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT AAC Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser Met Ser Asn 1735 1740 1745	5469
15	AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT CTC Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr Leu 1750 1760	5517
20	TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT GTT Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn Val 1765 1770 1775	5565
25	GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys 1780 1785 1790 1795	5613
	GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu 1800 1805 1810	5661
30	GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT Glu Leu Val Thr Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile 1815 1820 1825	5709
35	AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala 1830 1835 1840	5757
40	TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glu Thr Ile 1845 1850 1855	5805
45	AAA AAA GTG AAA GAC ATA TTT ACA GAC AGT TTC AGT AAA GTA ATT AAG Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys 1860 1865 1870 1875	5853
	GAA AAC AAC GAG AAT AAA TCA AAA ATT TGC CAA ACG AAA ATT ATG GCA Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys Ile Met Ala 1880 1885 1890	5901
50	GGT TGT TAC GAG GCA TTG GAT GAT TCA GAG GAT ATT CTT CAT AAC TCT Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu His Asn Ser 1895 1900 1905	5949
55	CTA GAT AAT GAT GAA TGT AGC ACG CAT TCA CAT AAG GTT TTT GCT GAC Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val Phe Ala Asp 1910 1915 1920	5997
60	ATT CAG AGT GAA GAA ATT TTA CAA CAT AAC CAA AAT ATG TCT GGA TTG Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met Ser Gly Leu 1925 1930 1935	6045
	GAG AAA GTT TCT AAA ATA TCA CCT TGT GAT GTT AGT TTG GAA ACT TCA	6093

	Glu 1940	Lys	Val	Ser	_	Ile 1945	Ser	Pro	Cys		Val 1950	Ser	Leu	Glu	Thr 1	Ser .955	
5				Lys					Lys					Val	TCA Ser L970		6141
10			Thr					Ser					Lys		GTC Val		6189
15		Ser					Gln					Val			GAA Glu		6237
20	Glu					Gln					Val				AGT Ser		6285
20					Gln					Glu					CGT Arg		6333
25				Leu					Gly					Val	GTA Val 2050		6381
30			Ala					Ser					Lys		GTT Val		6429
35		Leu					His					Val			GAA Glu		6477
40	Asp					Glu					Tyr				TCT Ser		6525
40					Lys					Val					CCA Pro		6573
45				Asn					Lys					Glu	TTT Phe 2130		6621
50			Asn					Glu					Glu		AAT Asn		6669
55		Ile					Tyr					Gln			AAA Lys		6717
60	Gln					Thr					Val				CAT His		6765
															ATT Ile		6813

5	AAC A			Gly					Asp					Ile		7581
10	AAT ( Asn (		Ile					Lys					Gln			7629
10	GTA . Val	Thr					Glu					Asp				7677
15	CTT Leu 2					Asp					Arg					7725
20	AGG Arg 2500				Phe					Ser					Lys	7773
25	TCC . Ser			Pro					Lys					Gly		7821
30	CCC Pro		Ala					Gln					Gly			7869
30	CAT His	Cys					Ser					Ser				7917
35	ACT Thr 2					Gly					${\tt Trp}$					7965
40	CAG Gln 2580				Gly					Pro					Lys	8013
45	GGA Gly			Glu					Leu	TGT Cys 2605				Gly		8061
50	CCA Pro		Leu					Trp					Tyr			8109
		Trp					Met			GCC Ala		Pro				8157
55	Asn					Pro				CTT Leu	Leu					8205
60					Ile					AGA Arg					Lys	8253

5				Asp					Lys					Cys	GTT Val 2690		8301
3			Ile					Asn					Ser		AAT Asn		8349
10		Ser					Gln					Ile			ACA Thr		8397
15	Gly					Lys					Pro				GCT Ala		8445
20					Arg					Gln					CAT His		8493
٦٢				Val					Ala					Glu	GCC Ala 2770		8541
25			Leu					Ser					Arg		GCT Ala		8589
30		Tyr					Phe					Arg			CCT Pro		8637
35	Pro					Phe					Asn				GTT Val		8685
40					Arg					Gln					ACA Thr		8733
45				Tyr					Glu					Lys	GAA Glu 2850		8781
43			Tyr					Gln					Ala		TTC Phe		8829
50		Ile					Glu					Asn			AAA Lys		8877
55	Tyr					Ala					Gln				TTG Leu		8925
60					Leu					Lys					CCA Pro		8973
	TAC	CTT	GAG	GGT	TAT	TTC	AGT	GAA	GAG	CAG	TTA	AGA	GCC	TTG	AAT	AAT	9021

	Tyr	Leu	Glu	_	Tyr 2920	Phe	Ser	Glu		Gln 2925	Leu	Arg	Ala		Asn 2930	Asn	
5		AGG Arg	Gln					Lys					Ile				9069
10		AGG Arg					Ser					Glu					9117
15	Arg	GAT Asp 2965				Val					Ile						9165
20		GAA Glu			Ser					Ile					Ser		9213
		TAT Tyr		Leu					Lys					Tyr			9261
25		ACT Thr	Ser					Lys					Asn				9309
30		GCG Ala					Gln					Pro					9357
35	Ile	TTA Leu 3045				Tyr					Pro						9405
40		TTA Leu			Asp					Cys					Leu		9453
		TTT Phe		Val					Lys					Pro			9501
45		TTG Leu	Ser					Asn					Lys				9549
50		CTT					Ile					Leu					9597
55	Asn	CTC Leu 3125				Pro					Gly						9645
60		GGA Gly			Ser					Ser					His		9693
		GAG Glu															9741

3160 3165 3170

5	Leu Cys Asn			ATG CAT ATA CTG Met His Ile Leu	9789
10				GAC TGT ACT TCA Asp Cys Thr Ser 3200	9837
15		Ile Ile Pro (		AAC AAG CTT CTG Asn Lys Leu Leu 3215	9885
				CCT TTA TCA CTT Pro Leu Ser Leu 3230	9933
20			Thr Pro Val	TCA GCC CAG ATG Ser Ala Gln Met 245	9981
25	Ser Cys Lys			GAC CAA AAG AAC Asp Gln Lys Asn	10029
30				CTG CCT TTA CCT Leu Pro Leu Pro 3280	10077
35		Cys Thr Phe		GCT GCA CAG AAG Ala Ala Gln Lys 3295	10125
33				GAA ACA CCC ATA Glu Thr Pro Ile 3310	10173
40			Met Thr Pro	TTT AAA AAA TTC Phe Lys Lys Phe 325	10221
45	Ser Leu Leu			GAC GAA GAA CTT Asp Glu Glu Leu	10269
50				ACA GGA GAA AAA Thr Gly Glu Lys 3360	10317
55		Glu Ser Thr		CCC ACC AGT TCA Pro Thr Ser Ser 3375	10365
22				TCT CTG ATC AAA Ser Leu Ile Lys 3390	10413
60			Glu Glu Cys	GAG AAA AAT AAG Glu Lys Asn Lys 8405	10461

## ATT ACA ACT AAA AAA TAT ATC TAA Ile Thr Thr Lys Lys Tyr Ile

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 3418 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein (v) FRAGMENT TYPE: internal
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe Glu Glu Leu Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu Glu Ser Glu His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr

Pro Gln Arg Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile Phe Lys Glu Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys

Glu Leu Asp Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser

Arg His Lys Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp 

Asp Val Ser Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val 

Val Leu Gln Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val 

Cys Gly Ser Leu Phe His Thr Pro Lys Phe Val Lys Gly Arg Gln Thr 

Pro Lys His Ile Ser Glu Ser Leu Gly Ala Glu Val Asp Pro Asp Met 

Ser Trp Ser Ser Leu Ala Thr Pro Pro Thr Leu Ser Ser Thr Val 

Leu Ile Val Arg Asn Glu Glu Ala Ser Glu Thr Val Phe Pro His Asp 

Thr Thr Ala Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu 

Lys Lys Asn Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr 

Asn Gln Arg Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn 

Ser Phe Lys Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro 

Asn Val Leu Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu 

Glu Asp Ser Phe Ser Leu Cys Phe Ser Lys Cys Arg Thr Lys Asn Leu 

Gln Lys Val Arg Thr Ser Lys Thr Arg Lys Lys Ile Phe His Glu Ala

	Asn	Ala	Asp	Glu 340	Cys	Glu	Lys	Ser	Lys 345	Asn	Gln	Val	Lys	Glu 350	Lys	Tyr
5	Ser	Phe	Val 355		Glu	Val	Glu	Pro 360		Asp	Thr	Asp	Pro 365		Asp	Ser
•	Asn	Val		His	Gln	Lys	Pro 375		Glu	Ser	Gly	Ser 380		Lys	Ile	Ser
	Lys 385	Glu	Val	Val	Pro	Ser 390	Leu	Ala	Cys	Glu	Trp 395	Ser	Gln	Leu	Thr	Leu 400
10	Ser	Gly	Leu	Asn	Gly 405	Ala	Gln	Met	Glu	Lys 410	Ile	Pro	Leu	Leu	His 415	Ile
	Ser	Ser	Cys	Asp 420	Gln	Asn	Ile	Ser	Glu 425	Lys	Asp	Leu	Leu	Asp 430	Thr	Glu
15	Asn	Lys	Arg 435	Lys	Lys	Asp	Phe	Leu 440	Thr	Ser	Glu	Asn	Ser 445	Leu	Pro	Arg
		Ser 450				_	455		_			460				
	465	Asn	_		_	470					475				_	480
20		Leu			485					490					495	
		Phe		500		_	_		505		_		_	510		
25		Glu	515					520		_			525			
		Lys 530	_				535				_	540				
30	545	Cys				550					555					560
30	_	Ser Gly	_		565					570					575	
		Ile		580					585	_			_	590		_
35		Lys	595					600				_	605		_	_
		610 Glu					615					620				
40	625	Ser				630					635		_			640
		Ser		_	645		_			650	_				655	
	Asn	Glu	Thr	660 Cys	Ser	Asn	Asn	Thr	665 Val	Ile	Ser	Gln	Asp	670 Leu	Asp	Tyr
45	Lys	Glu	675 Ala	Lys	Cys	Asn	_	680 Glu	Lys	Leu	Gln		685 Phe	Ile	Thr	Pro
		690 Ala	Asp	Ser	Leu		695 Cys	Leu	Gln	Glu	_	700 Gln	Cys	Glu	Asn	-
50	705 Pro	Lys	Ser	Lys	-	710 Val	Ser	Asp	Ile	-	715 Glu	Glu	Val	Leu		720 Ala
	Ala	Cys	His	Pro 740	725 Val	Gln	His	Ser	Lys 745	730 Val	Glu	Tyr	Ser	Asp 750	735 Thr	Asp
55	Phe	Gln	Ser 755		Lys	Ser	Leu	Leu 760		Asp	His	Glu	Asn 765		Ser	Thr
	Leu	Ile 770		Thr	Pro	Thr	Ser 775		Asp	Val	Leu	Ser 780		Leu	Val	Met
	Ile 785	Ser	Arg	Gly	Lys	Glu 790		Tyr	Lys	Met	Ser 795		Lys	Leu	Lys	Gly 800
60	Asn	Asn	Tyr	Glu	Ser 805	Asp	Val	Glu	Leu	Thr 810	Lys	Asn	Ile	Pro	Met 815	Glu
	Lys	Asn	Gln	Asp	Val	Cys	Ala	Leu	Asn	Glu	Asn	Tyr	Lys	Asn	۷al	Glu

				820					825					830		
	Leu	Leu	Pro 835	Pro	Glu	Lys	Tyr	Met 840	Arg	Val	Ala	Ser	Pro 845	Ser	Arg	Lys
5	Val	Gln 850	Phe	Asn	Gln	Asn	Thr 855	Asn	Leu	Arg	Val	Ile 860	Gln	Lys	Asn	Gln
	865			Thr		870		_			875			_		880
10				Ser	885					890					895	
				Asn 900					905					910		
	_		915	Cys				920			_		925			
15		930	_	Asp		_	935	_				940				_
	945	_		Val	_	950					955	_				960
20				Lys	965			-		970		_		_	975	
				Asp 980	_				985					990		
25	_		995	Leu		_		1000	)				1005	5		
25		1010	)	Ala			1015	5		_		1020	)			
	1025	5		Lys		1030	)	_	_		1035	5		_		104
30				Cys	1045	5				1050	)			_	1055	5
	-	-		Ser 1060	) -		GIII	ser	1065		TIIL	vai	ser	1070		neu
				1/21	77 - 7	7727	Car	7 cn	Circ	TTC	Λcn	Car	uic	Tla	Thr	Dro
2.5			1075	5			Ser	1080	)				1085	5		
35	Gln	Met 1090	1075 Leu )	Phe	Ser	Lys	Gln 1095	1080 Asp	) Phe	Asn	Ser	Asn 1100	1089 His	Asn	Leu	Thr
35	Gln Pro 1109	Met 1090 Ser	1075 Leu ) Gln	Phe Lys	Ser Ala	Lys Glu 1110	Gln 1095 Ile	1080 Asp Thr	Phe Glu	Asn Leu	Ser Ser	Asn 1100 Thr	1089 His ) Ile	Asn Leu	Leu Glu	Thr Glu 112
35	Gln Pro 1105 Ser	Met 1090 Ser Gly	1079 Leu ) Gln Ser	Phe Lys Gln	Ser Ala Phe	Lys Glu 1110 Glu	Gln 1095 Ile ) Phe	1080 Asp Thr	Phe Glu Gln	Asn Leu Phe 1130	Ser Ser 1115 Arg	Asn 1100 Thr Xaa	1089 His ) Ile Pro	Asn Leu Ser	Leu Glu Tyr 1135	Thr Glu 112 Ile
	Gln Pro 1105 Ser Leu	Met 1090 Ser Gly Gln	1075 Leu ) Gln Ser Lys	Phe Lys Gln Ser	Ser Ala Phe 1129 Thr	Lys Glu 1110 Glu Phe	Gln 1095 Ile ) Phe Glu	1080 Asp Thr Thr	Phe Glu Gln Pro 1145	Asn Leu Phe 1130 Glu	Ser Ser 1115 Arg ) Asn	Asn 1100 Thr Xaa Gln	1085 His ) Ile Pro Met	Asn Leu Ser Thr	Leu Glu Tyr 1135 Ile	Thr Glu 112 Ile Leu
40	Gln Pro 1109 Ser Leu Lys	Met 1090 Ser Gly Gln Thr	1075 Leu Gln Ser Lys Thr	Phe Lys Gln Ser 1140 Ser	Ser Ala Phe 1129 Thr Glu	Lys Glu 1110 Glu Phe Glu	Gln 1095 Ile Phe Glu Cys	1080 Asp Thr Thr Val Arg	Phe Glu Gln Pro 1145 Asp	Asn Leu Phe 1130 Glu Ala	Ser Ser 1115 Arg Arg Asn Asp	Asn 1100 Thr Xaa Gln Leu	1089 His Ile Pro Met His 1169	Asn Leu Ser Thr 1150 Val	Leu Glu Tyr 1135 Ile	Thr Glu 112 Ile Leu Met
	Gln Pro 1109 Ser Leu Lys Asn	Met 1090 Ser Gly Gln Thr	1075 Leu Gln Ser Lys Thr 1155 Pro	Phe Lys Gln Ser 1140 Ser Ser	Ser Ala Phe 1125 Thr Glu Ile	Lys Glu 1110 Glu Phe Glu Glu	Gln 1095 Ile Phe Glu Cys Gln 1175	1080 Asp Thr Thr Val Arg 1160 Val	Phe Glu Gln Pro 1149 Asp Asp	Asn Leu Phe 1130 Glu Ala Ser	Ser Ser 1115 Arg Asn Asp Ser	Asn 1100 Thr Xaa Gln Leu Lys 1180	1089 His Ile Pro Met His 1169 Gln	Asn Leu Ser Thr 1150 Val Phe	Leu Glu Tyr 1135 Ile Ile Glu	Thr Glu 112 Ile Leu Met Gly
40	Gln Pro 1109 Ser Leu Lys Asn Thr 1189	Met 1090 Ser 5 Gly Gln Thr Ala 1170 Val	1075 Leu Gln Ser Lys Thr 1155 Pro	Phe Lys Gln Ser 1140 Ser Ser Ile	Ser Ala Phe 1125 Thr Glu Ile Lys	Lys Glu 1110 Glu Phe Glu Gly Arg 1190	Gln 1095 Ile Phe Glu Cys Gln 1175 Lys	1080 Asp Thr Thr Val Arg 1160 Val	Phe Glu Gln Pro 1145 Asp Asp Asp	Asn Leu Phe 1130 Glu Ala Ser Gly	Ser Ser 1115 Arg Asn Asp Ser Leu 1195	Asn 1100 Thr Xaa Gln Leu Lys 1180 Leu	1085 His Dile Pro Met His 1165 Gln Lys	Asn Leu Ser Thr 1150 Val Phe Asn	Leu Glu Tyr 1135 Ile Ile Glu Asp	Thr Glu 112 Ile Leu Met Gly Cys 120
40	Gln Pro 1109 Ser Leu Lys Asn Thr 1189 Asn	Met 1090 Ser 5 Gly Gln Thr Ala 1170 Val	1075 Leu Cln Ser Lys Thr 1155 Pro Clu Ser	Phe Lys Gln Ser 1140 Ser Ser Ile Ala	Ser Ala Phe 1125 Thr Glu Ile Lys Ser 1205	Lys Glu 1110 Glu Phe Glu Gly Arg 1190 Gly	Gln 1099 Ile Phe Glu Cys Gln 1179 Lys Tyr	1080 Asp Thr Thr Val Arg 1160 Val Phe Leu	Phe Glu Gln Pro 1145 Asp Asp Ala Thr	Asn Leu Phe 1130 Glu Ala Ser Gly Asp 1210	Ser Ser 1115 Arg Asn Asp Ser Leu 1195 Glu	Asn 1100 Thr Xaa Gln Leu Lys 1180 Leu Asn	IO85 His The Pro Met His 1165 Gln Lys Glu	Asn Leu Ser Thr 1150 Val Phe Asn	Leu Glu Tyr 1135 Ile Glu Asp Gly 1215	Thr Glu 112 Ile Leu Met Gly Cys 120 Phe
40 45	Gln Pro 1109 Ser Leu Lys Asn Thr 1189 Asn	Met 1090 Ser 5 Gly Gln Thr Ala 1170 Val 5 Lys	1075 Leu Cln Ser Lys Thr 1155 Pro Clu Ser Phe	Phe Lys Gln Ser 1140 Ser Ser Ile Ala Tyr 1220	Ser Ala Phe 1125 Thr Glu Ile Lys Ser 1205 Ser	Clu 1110 Glu Phe Glu Gly Arg 1190 Gly Ala	Gln 1099 Ile Phe Glu Cys Gln 1179 Lys Tyr His	1080 Asp Thr Thr Val Arg 1160 Val Phe Leu Gly	Phe Glu Gln Pro 1145 Asp Asp Ala Thr	Asn Leu Phe 1130 Glu Ala Ser Gly Asp 1210 Lys	Ser Ser 1115 Arg Asn Asp Ser Leu 1195 Glu Leu	Asn 1100 Thr Xaa Gln Leu Lys 1180 Leu Asn	1085 His Ile Pro Met His 1165 Gln Lys Glu Val	Asn Leu Ser Thr 1150 Val Phe Asn Val Ser 1230	Leu Glu Tyr 1135 Ile Glu Asp Gly 1215 Thr	Thr  Glu 112 Ile Leu Met Gly Cys 120 Phe Glu
40 45 50	Gln Pro 1109 Ser Leu Lys Asn Thr 1189 Asn Arg	Met 1090 Ser Gly Gln Thr Ala 1170 Val Lys Gly Leu	1075 Leu Cln Ser Lys Thr 1155 Pro Clu Ser Phe Gln 1235	Phe Lys Gln Ser 1140 Ser Ile Ala Tyr 1220 Lys	Ser Ala Phe 1125 Thr Glu Ile Lys Ser 1205 Ser Ala	Clu Glu Phe Glu Gly Arg 1190 Gly Ala Val	Gln 1099 Ile Phe Glu Cys Gln 1179 Lys Tyr His	1080 Asp Thr Thr Val Arg 1160 Val Phe Leu Gly Leu 1240	Phe Glu Gln Pro 1145 Asp Asp Ala Thr 1225 Phe	Asn Leu Phe 1130 Glu Ala Ser Gly Asp 1210 Lys Ser	Ser Ser 1115 Arg Asn Asp Ser Leu 1195 Glu Leu Asp	Asn 1100 Thr Xaa Gln Leu Lys 1180 Leu Asn Asn	IO85 His Ile Pro Met His 1165 Gln Lys Glu Val Glu 1245	Asn Leu Ser Thr 1150 Val Phe Asn Val Ser 1230 Asn	Leu Glu Tyr 1135 Ile Glu Asp Gly 1215 Thr	Thr  Glu 112 Ile Leu Met Gly Cys 120 Phe Glu Ser
40 45	Gln Pro 1109 Ser Leu Lys Asn Thr 1189 Asn Arg Ala Glu	Met 1090 Ser Gly Gln Thr Ala 1170 Val Lys Gly Leu Glu 1250	1075 Leu Gln Ser Lys Thr 1155 Pro Glu Ser Phe Gln 1235 Thr	Phe Lys Gln Ser 1140 Ser Ile Ala Tyr 1220 Lys Ser	Ala Phe 1125 Thr Glu Ile Lys Ser 1205 Ser Ala Ala	Clu 1110 Glu Phe Glu Gly Arg 1190 Gly Ala Val	Gln 1095 Ile Phe Glu Cys Gln 1175 Lys Tyr His Lys Val 1255	1080 Asp Thr Thr Val Arg 1160 Val Phe Leu Gly Leu 1240 His	Phe Glu Gln Pro 1145 Asp Asp Ala Thr Thr 1225 Phe	Asn Leu Phe 1130 Glu Ala Ser Gly Asp 1210 Lys Ser Ile	Ser Ser 1115 Arg Asn Asp Ser Leu 1195 Glu Leu Asp Ser	Asn 1100 Thr Xaa Gln Leu Lys 1180 Leu Asn Asn Ile Leu 1260	IO85 His Ile Pro Met His 1165 Gln Lys Glu Val Glu 1245 Ser	Asn Leu Ser Thr 1150 Val Phe Asn Val Ser 1230 Asn Ser	Leu Glu Tyr 1135 Ile Glu Asp Gly 1215 Thr Ile Ser	Thr  Glu 112 Ile Leu Met Gly Cys 120 Phe Glu Ser Lys
40 45 50	Gln Pro 1109 Ser Leu Lys Asn Thr 1189 Asn Arg Ala Glu Cys 1269	Met 1090 Ser Gly Gln Thr Ala 1170 Val 5 Lys Gly Leu Glu 1250 His	1075 Leu Gln Ser Lys Thr 1155 Pro Glu Ser Phe Gln 1235 Thr	Phe Lys Gln Ser 1140 Ser Ser Ile Ala Tyr 1220 Lys Ser Ser	Ala Phe 1125 Thr Glu Ile Lys Ser 1205 Ser Ala Ala Val	Clu 1110 Glu Fhe Glu Gly Arg 1190 Gly Ala Val Glu Val 1270	Gln 1095 Ile Phe Glu Cys Gln 1175 Lys Tyr His Lys Val 1255 Ser	1080 Asp Thr Thr Val Arg 1160 Val Phe Leu Gly Leu 1240 His Met	Phe Glu Gln Pro 1145 Asp Asp Ala Thr Thr 1225 Phe Pro	Asn Leu Phe 1130 Glu Ala Ser Gly Asp 1210 Lys Ser Ile Lys	Ser Ser 1115 Arg Asn Asp Ser Leu 1195 Glu Leu Asp Ser Ile 1275	Asn 1100 Thr Xaa Gln Leu Lys 1180 Leu Asn Asn Ile Leu 1260 Glu	IO89 His Fro Met His 1169 Gln Lys Glu Val Glu 1249 Ser Asn	Asn Leu Ser Thr 1150 Val Phe Asn Val Ser 1230 Asn Ser His	Leu Glu Tyr 1135 Ile Glu Asp Gly 1215 Thr Ile Ser Asn	Thr Glu 112 Ile Leu Met Gly Cys 120 Phe Glu Ser Lys Asp 128
40 45 50	Gln Pro 1109 Ser Leu Lys Asn Thr 1189 Asn Arg Ala Glu Cys 1269 Lys	Met 1090 Ser Gly Gln Thr Ala 1170 Val 5 Lys Gly Leu Glu 1250 His	IO75 Leu Gln Ser Lys Thr 1155 Pro Glu Ser Phe Gln 1235 Thr Asp	Phe Lys Gln Ser 1140 Ser Ile Ala Tyr 1220 Lys Ser	Ser Ala Phe 1125 Thr Glu Ile Lys Ser 1205 Ser Ala Ala Val Glu 1285	Clu Clu Clu Clu Clu Clu Cly Cly Cly Cly Clu Clu Clu Cly	Gln 1099 Ile Phe Glu Cys Gln 1179 Lys Tyr His Lys Val 1259 Ser	1080 Asp Thr Thr Val Arg 1160 Val Phe Leu Gly Leu 1240 His Met Asn	Phe Glu Gln Pro 1145 Asp Asp Ala Thr Thr 1225 Phe Pro Phe Lys	Asn Leu Phe 1130 Glu Ala Ser Gly Asp 1210 Lys Ser Ile Lys Cys 1290	Ser Ser 1119 Asn Asp Ser Leu 1199 Glu Leu Asp Ser Ile 1279 Gln	Asn 1100 Thr Xaa Gln Leu Lys 1180 Leu Asn Ale Leu 1260 Glu Leu	1089 His Ile Pro Met His 1169 Gln Lys Glu Val Glu 1249 Ser Asn	Asn Leu Ser Thr 1150 Val Phe Asn Val Ser 1230 Asn Ser His	Leu Glu Tyr 1135 Ile Glu Asp Gly 1215 Thr Ile Ser Asn Gln 1295	Thr Glu 112 Ile Leu Met Gly Cys 120 Phe Glu Ser Lys Asp 128 Asn

	Tyr	Lys	Arg 1315		Thr	Glu	Asn	Glu 1320		Asn	Lys	Tyr	Thr 1325		Ala	Ser
5	_	Asn 1330	)				1335	5				1340	)			
	1345			-		1350	)	_			1355	5				136
		His			1365	5				1370	)			_	1375	5
10		Thr		1380	)				1385	5				1390	)	
		Lys	1395	5			_	1400	) -				1405	5		
15		Thr 1410	)				1415	5				1420	)			
	1425	Thr	Pne	Pne	GIII	1430		ser	GIÀ	гуѕ	1435		ser	vai	Ald	LуS 144
		Ser	Phe	Asn	Lys			Asn	Phe	Phe			Lys	Pro	Glu	
					1445					1450			_		1455	
20		His		1460	)				1465	5				1470	)	
		Lys	1475	5				1480	)			_	1485	5	-	
25	Lys	Ile 1490		Lys	GIu	Ser	Val 1495		Val	Gly	Thr	GIY 1500		GIn	Leu	Val
23	Thr 1509	Phe		Gly	Gln	Pro 1510	Glu		Asp	Glu	Lys 1515	Ile		Glu	Pro	Thr 152
	Leu	Leu	Gly	Phe	His 1525		Ala	Ser	Gly	Lys 1530		Val	Lys	Ile	Ala 1535	
30		Ser		1540	)				1545	5				1550	)	
		Ser	1555	5				1560	)				1565	5		_
35		Arg 1570	)				1575	5				1580	)			
	1589	Thr	Ala	Ala	Pro	ьув 1590		ьуs	GIU	мет	1595		ser	ьeu	Asn	160
		Lys	Asn	Leu	Val 1605	Ser		Glu	Thr	Val 1610	Val		Pro	Lys	Leu 1615	Leu
40		Asp		1620	)				1625	5				1630	)	
			1635	5		_		1640	)				1645	5		
45		Ser 1650	)			_	1655	5				1660	)			
	1665	Asn	ser	Ala	ьeu	1670		Tyr	Thr	ser	1675		Arg	ьуs	Thr	ser 168
		Ser	Gln	Thr	Ser 1685	Leu		Glu	Ala	Lys 1690	Lys		Leu	Arg	Glu 1695	Gly
50	Ile	Phe	Asp	Gly 1700	Gln		Glu	Arg	Ile 1705	Asn		Ala	Asp	Tyr 1710	Val	
		Tyr	1715	5				1720	)				1725	5		_
55		Asn 1730	)				1735	5	_		_	1740	)			
	Met 1749	Ser	Asn	ser	Tyr	Ser 1750		Hls	ser	Asp	Glu 1755		Tyr	Asn	Asp	Ser 176
•		Tyr	Leu	Ser	Lys 1765	Asn		Leu	Asp	Ser 1770	Gly		Glu	Pro	Val 1775	Leu
60	Lys	Asn	Val		Asp		Lys	Asn		Ser		Ser	Lys		Ile	
	Asn	Val	Lys	1780 Asp		Asn	Ala	Tyr	1789 Pro		Thr	Val	Asn	1790 Glu		Ile

		1	L795	5				1800	)				1805	5		
	Cys V	al G 810	3lu	Glu	Leu	Val	Thr 1815	Ser		Ser	Pro	Cys 1820	Lys		Lys	Asn
5	Ala A 1825	la I	lle	Lys	Leu	Ser 1830		Ser	Asn	Ser	Asn 1835		Phe	Glu	Val	Gly 184
	Pro P	ro A	Ala	Phe	Arg 1845	Ile		Ser	Gly	Lys 1850	Ile		Cys	Val	Ser 1855	His
10	Glu T	hr I	lle	Lys 1860		Val	Lys	Asp	Ile 1865		Thr	Asp	Ser	Phe 1870		Lys
	Val I		.ys .875		Asn	Asn	Glu	Asn 1880	Lys		Lys	Ile	Cys 1885	Gln		Lys
	Ile Me	et A 890	la	Gly	Cys	Tyr	Glu 1895	Ala		Asp	Asp	Ser 1900	Glu		Ile	Leu
15	His A	sn S	Ser	Leu	Asp			Glu	Cys	Ser			Ser	His	Lys	
	1905 Phe A	la A	Asp	Ile				Glu	Ile				Asn	Gln		
	Ser G	lv L	eu	Glu	1925 Lvs		Ser	Lvs	Tle	1930 Ser		Cvs	Asn	Val	1935 Ser	
20				1940	)			_	1945	5		-	-	1950	)	
	Glu Tl	1	.955					1960	)				1965	5		
		970					1975	;				1980	)		_	_
25	Ser Va 1985	al G	ln	Val	Ser	Asp 1990		Ser	Leu	Gln	Asn 1995		Arg	Gln	Val	Phe 200
	Ser G	lu I	le	Glu	Asp 2005		Thr	Lys	Gln	Val 2010		Ser	Lys	Val	Leu 2015	
2.0	Lys Se	er A	sn			Ser	Asp	Gln			Arg	Glu	Glu			Ala
30	Ile A		hr :035			His	Leu	Ile 2040			Lys	Gly	Phe 2045			Asn
	Val Va				Ser	Ala	Phe 2055	Ser		Phe	Ser	Thr 2060	Ala		Gly	Lys
35	Gln Va 2065		er	Ile	Leu	Glu 2070	Ser		Leu	His	Lys 2075	Val		Gly	Val	Leu 208
	Glu G	lu P	he	Asp		Ile		Thr	Glu	His			His	Tyr	Ser	
	Thr Se	er A	ırg	Gln	2085 Asn		Ser	Lys	Ile	2090 Leu		Ara	Val	Asp	2095 Lvs	
40	Asn Pi			2100	1				2105	5				2110	)	_
	Glu Pl	2	115					2120	)				2125	i		_
4.5	2	130					2135					2140	_			
45	Asn As 2145					2150	)				2155					216
	Asp Ly	ys G	ln	Gln	Leu 2165		Leu	Gly	Thr	Lys 2170		Ser	Leu	Val	Glu 2175	
50	Ile H	is V		Leu 2180		Lys	Glu	Gln	Ala 2185		Pro	Lys	Asn	Val 2190	Lys	
	Glu I			Lys		Glu	Thr	Phe 2200	Ser		Val		Val 2205	Lys		Asn
	Ile G	lu V 210	al	Cys	Ser	Thr	Tyr 2215	Ser		Asp			Asn		Phe	Glu
55	Thr G1 2225		.la	Val	Glu	Ile 2230	Ala		Ala			Glu		Asp	Glu	Leu 224
	Thr As	sp S	er	Lys	Leu 2245	Pro		His	Ala		His		Leu	Phe	Thr 2255	Cys
60	Pro G	lu A		Glu 2260	Glu		Val	Leu	Ser 2265	Asn		Arg	Ile	Gly 2270	Lys	
	Arg G	ly G 2		Pro		Ile		Val 2280	Gly		Pro		Ile 2285	Lys		Asn

	Leu	Leu 2290	Asn O	Glu	Phe	Asp	Arg 229		Ile	Glu	Asn	Gln 2300		Lys	Ser	Leu
	Lys	Ala	Ser	Lvs	Ser	Thr	Pro	Asp	Glv	Thr	Ile	Lvs	Asp	Ara	Ara	Let
5	2305			4		2310			1		2315		F	3	5	232
			His	His	Val 2325	Ser		Glu	Pro	Ile 2330	Thr		Val	Pro	Phe 2335	Arg
	The	The	Tira	<i>α</i> 1			<b>a</b> 1	T1.	<b>01</b> -			7	Dh.	ml		
	1111	TILL	Lys			GIII	GIU	TTE			Pro	ASI	Pne			Pro
10				2340					2345			_		2350		
10			Glu 2355	5				2360	)				2365	5		
	Glu	Lys 2370	Ser	Ser	Ser	Asn	Leu 2375		Val	Ser	Gly	His 2380		Phe	Tyr	Glr
	Val		Ala	Thr	Ara	Agn			Met	Δra	Hie			Thr	Thr	G1s
15	2385				5	2390					2395				* ***	240
			Thr	Laze	17a l			Dro	Dro				Larg	802	ніс	
	AT 9	FIO	1111	цуз			vai	PIO	PIO			TILL	пув	ser		
	*** -	3	77- 7	<b>~</b> 1	2409			_	_	2410		_			2419	
	HIS	Arg	Val			Cys	vai				Asn	Leu	Glu			Arg
				2420					2425					2430		
20	Gln	Lys	Gln	Asn	Ile	Asp	Gly	His	Gly	Ser	Asp				Asn	Lys
			2435					2440					2445	5		
	Ile	Asn	Asp	Asn	Glu	Ile	His	Gln	Phe	Asn	Lys	Asn	Asn	Ser	Asn	Gln
		2450					2455				-	2460				
	Ala	Ala	Ala	Val	Thr	Phe	Thr	Lvs	Cvs	Glu	Glu			Len	Asn	Len
25	2465					2470		- 7 -	0,2	010	2475		110	cu	11Dp	248
			Ser	T.011	Gln			Λrα	7 an	Tlo			Mot	7 ~~	тіс	
	116	1111	SET	пец	2485		нта	Arg	Asp			Asp	Met	Arg		_
	_	_	~ 7	_					_	2490					2495	
	ьуs	гуз	Gln			Arg	Val	Phe			Pro	Gly	Ser	Leu	Tyr	Leu
				2500					2505					2510		
30	Ala	Lys	Thr	Ser	Thr	Leu	Pro	Arg	Ile	Ser	Leu	Lys	Ala	Ala	Val	Gly
			2515	5				2520	)				2525	5		
	Gly	Gln	Val	Pro	Ser	Ala	Cys	Ser	His	Lys	Gln	Leu	Tyr	Thr	Tyr	Gly
		2530					2535					2540	_		-	-
	Val	Ser	Lys	His	Cys	Ile	Lys	Ile	Asn	Ser	Lvs	Asn	Ala	Glu	Ser	Phe
35	2545		-		•	2550					2555					256
			His	Thr	Glu			Phe	Glv				T.e.11	Trn	Thr	
					2565		- y -	1110	O L y	2570		OCI	пси	TLD	2575	_
	T 3.70	C1.,	Tlo	~1 n			7	<b>a</b> 1	Q1			T1-	D	0		
	пуѕ	GLY	Ile			Ala	Asp				Leu	TIE	Pro			Asp
4.0		_		2580					2585		_			2590		
40	GIA	ьуs	Ala		гÀг	GIu	Glu			Arg	Ala	Leu			Thr	Pro
	_	_	2595					2600					2605			
	GIY	Val	Asp	Pro	Lys	Leu	Ile	Ser	Arg	Ile	Trp	Val	Tyr	Asn	His	Tyr
		2610						5								
	Arg	Trp	Ile	Ile	Trp	Lys	Leu	Ala	Ala	Met	Glu	Cys	Ala	Phe	Pro	Lys
45	2625	5				2630	)				2635					264
	Glu	Phe	Ala	Asn	Arg	Cys	Leu	Ser	Pro	Glu	Arq	Val	Leu	Leu	Gln	Leu
					2645					2650					2655	
	Lvs	Tvr	Arg	Tvr			Glu	Tle	Asp			Ara	Δra	Ser		
	-1-	- 1 -	5	2660		****	CIG	110	2665		CCI	nr 9	AL 9	2670		110
50	Laze	Lare	Tla			7 ~~	7 an	7 00			70 7 -	T	TT la sa			T
50	цуз	цуз	Ile		GIU	Arg	Asp			Ala	Ата	гуя			Val	ьeu
	<b>G</b>	** - 7	2675		~ 7		_	2680			_	<b>-</b>	2685		_	
	Cys		Ser	Asp	Пе	TIE			Ser	Ala	Asn			Glu	Thr	Ser
		2690					2695					2700				
	Ser	Asn	Lys	Thr	Ser	Ser	Ala	Asp	Thr	Gln	Lys	Val	Ala	Ile	Ile	Glu
55	2705	i				2710	)				2715					272
	Leu	Thr	Asp	Gly	Trp	Tyr	Ala	Val	Lys	Ala	Gln	Leu	Asp	Pro	Pro	Leu
				_	2725				-	2730			-		2735	
	Leu	Ala	Val	Leu			Glv	Ara	Leu			Glv	Gln	Lvs		
				2740			1	ر	2745			1		2750		_~~
60	Leu	His	Gly			Leu	Val	Glv			Asn	Δla	Cve			T.e.i
			2755					2760					2765			LCu
			Dro					-,00					2,00	•		

		2770	)				2775	5				2780	)			
	Pro 278		Arg	Trp	Tyr	Thr 2790	Lys O	Leu	Gly	Phe	Phe 2795		Asp	Pro	Arg	Pro 280
5	Phe	Pro	Leu	Pro	Leu 2805		Ser	Leu	Phe	Ser 2810		Gly	Gly	Asn	Val 281	_
				2820	)		Gln		2825	5				2830	)	
10			2835	5			Tyr	2840	)				2845	5		
		2850	)				Val 2855	5				2860	) _			
	2869	5				2870					2875	5				288
15					2885	5	Ser			2890	)	_			2895	5
				2900	)		Glu		2905	5			_	2910	)	
20			2915	5			Gly	2920	)				2925	5	_	
		2930	)				Met 2935	5		_	_	2940	)			
25	2945	5				2950					2955	5		_		296
25					2965	5	Thr			2970	)				2975	5
				2980	)		Asp		2985	5				2990	)	
30			2995	5			Leu	3000	)		_	_	3005	5	_	
		3010	)				Lys 3015	;				3020	)			
2.5	3025	5				3030					3035	5				304
35					3045	5	Gln			3050	)				3055	5
				3060	)		Pro		3065	5				3070	)	
40			3075	5			Val	3080	)				3085	5		
		3090	)				Asp 3095		_	-		3100	)			•
45	3105	5				3110					3115	;				312
40					3125	5	Trp			3130	)				3135	;
				3140	)		Phe		3145	;				3150	-	
50			3155	5			Phe	3160	1				3165	5		
		3170	)				Glu 3175					3180	ı			
55	3185	5				3190					3195	;		_		320
33					3205	;	Ile			3210	)				3215	j
				3220			Glu		3225	;				3230		
60			3235	i			Ser	3240	)				3245	;		
	1111	3250		sei	сув	пув	Gly 3255		пÀг	ĿΙU	тте	Asp 3260		GIN	ьys	Asn

	Cys Lys Lys Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro 3265 3270 328	
	3275 328 Pro Pro Val Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys	
5	3285 3290 3295	
	Ala Phe Gln Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile 3300 3305 3310	
	Lys Lys Lys Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe 3315 3320 3325	
10	Asn Glu Ile Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu 3330 3340	
	Ala Leu Ile Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys 3345 3350 336	
15	Gln Phe Ile Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser	
	Glu Asp Tyr Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys	
	Glu Gln Glu Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys	
20	3395 3400 3405 Gln Asp Thr Ile Thr Thr Lys Lys Tyr Ile 3410 3415	
25	(2) INFORMATION FOR SEQ ID NO:8:	
25	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 10485 base pairs</li></ul>	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: double (D) TOPOLOGY: linear	
30	(ii) MOLEGULE TUDE TOWN	
	(ii) MOLECULE TYPE: cDNA (ix) FEATURE:	
3 E	(A) NAME/KEY: Coding Sequence	
35	(B) LOCATION: 22910482 (D) OTHER INFORMATION: BRCA2 (OMI3)	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:	
40	GGTGGCGCGA GCTTCTGAAA CTAGGCGGCA GAGGCGGAGC CGCTGTGGCA CTGCTGCGCC	60
	TCTGCTGCGC CTCGGGTGTC TTTTGCGGCG GTGGGTCGCC GCCGGGAGAA GCGTGAGGGG ACAGATTTGT GACCGGCGCG GTTTTTGTCA GCTTACTCCG GCCAAAAAAG AACTGCACCT	120
		180 237
45	Met Pro Ile 1	
	GGA TCC AAA GAG AGG CCA ACA TTT TTT GAA ATT TTT AAG ACA CGC TGC Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys Thr Arg Cys	285
F.0	5 10 15	
50	AAC AAA GCA GAT TTA GGA CCA ATA AGT CTT AAT TGG TTT GAA GAA CTT	333
	Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe Glu Glu Leu	333
	25 30 35	
55	TCT TCA GAA GCT CCA CCC TAT AAT TCT GAA CCT GCA GAA GAA TCT GAA	381
	Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu Glu Ser Glu 40 45 50	
	30	
60	CAT AAA AAC AAC AAT TAC GAA CCA AAC CTA TTT AAA ACT CCA CAA AGG His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr Pro Gln Arg	429
	55 60 65	

5				TAT Tyr													477
J				ACT Thr													525
10				TTA Leu													573
15				ACA Thr													621
20				CTA Leu 135													669
25				GTA Val													717
				ACA Thr													765
30				AGT Ser													813
35				GCT Ala													861
40				GAA Glu 215													909
45	_		_	AGC Ser		_											957
				ATC Ile													1005
50				AGT Ser													1053
55				TGC Cys													1101
60				GTA Val 295													1149
	TTT	TCA	TTA	TGT	TTT	TCT	AAA	TGT	AGA	ACA	AAA	TAA	CTA	CAA	AAA	GTA	1197

	Phe	Ser	Leu 310	Cys	Phe	Ser	Lys	Cys 315	Arg	Thr	Lys	Asn	Leu 320	Gln	Lys	Val	
5				AAG Lys													1245
10				AAA Lys													1293
15				GAA Glu													1341
2.0				CCC Pro 375													1389
20				TTG Leu													1437
25				CAG Gln													1485
30				ATT Ile													1533
35				TTT Phe													1581
4.0				TCA Ser 455													1629
40				GAG Glu													1677
45				GCA Ala									_				1725
50				AAG Lys													1773
55				AGT Ser													1821
60				GCC Ala 535													1869
60				GAC Asp													1917

550 555 560

5	CCA Pro	GCC Ala 565	ACC Thr	ACC Thr	ACA Thr	CAG Gln	AAT Asn 570	TCT Ser	GTA Val	GCT Ala	TTG Leu	AAG Lys 575	AAT Asn	GCA Ala	GGT Gly	TTA Leu	1965
10	ATA Ile 580	TCC Ser	ACT Thr	TTG Leu	AAA Lys	AAG Lys 585	AAA Lys	ACA Thr	AAT Asn	AAG Lys	TTT Phe 590	ATT Ile	TAT Tyr	GCT Ala	ATA Ile	CAT His 595	2013
15	GAT Asp	GAA Glu	ACA Thr	TCT Ser	TAT Tyr 600	AAA Lys	GGA Gly	AAA Lys	AAA Lys	ATA Ile 605	CCG Pro	AAA Lys	GAC Asp	CAA Gln	AAA Lys 610	TCA Ser	2061
13	GAA Glu	CTA Leu	ATT Ile	AAC Asn 615	TGT Cys	TCA Ser	GCC Ala	CAG Gln	TTT Phe 620	GAA Glu	GCA Ala	AAT Asn	GCT Ala	TTT Phe 625	GAA Glu	GCA Ala	2109
20	CCA Pro	CTT Leu	ACA Thr 630	TTT Phe	GCA Ala	AAT Asn	GCT Ala	GAT Asp 635	TCA Ser	GGT Gly	TTA Leu	TTG Leu	CAT His 640	TCT Ser	TCT Ser	GTG Val	2157
25	AAA Lys	AGA Arg 645	AGC Ser	TGT Cys	TCA Ser	CAG Gln	AAT Asn 650	GAT Asp	TCT Ser	GAA Glu	GAA Glu	CCA Pro 655	ACT Thr	TTG Leu	TCC Ser	TTA Leu	2205
30	ACT Thr 660	AGC Ser	TCT Ser	TTT Phe	GGG Gly	ACA Thr 665	ATT Ile	CTG Leu	AGG Arg	AAA Lys	TGT Cys 670	TCT Ser	AGA Arg	AAT Asn	GAA Glu	ACA Thr 675	2253
2.5	TGT Cys	TCT Ser	AAT Asn	AAT Asn	ACA Thr 680	GTA Val	ATC Ile	TCT Ser	CAG Gln	GAT Asp 685	CTT Leu	GAT Asp	TAT Tyr	AAA Lys	GAA Glu 690	GCA Ala	2301
35	AAA Lys	TGT Cys	AAT Asn	AAG Lys 695	GAA Glu	AAA Lys	CTA Leu	CAG Gln	TTA Leu 700	TTT Phe	ATT Ile	ACC Thr	CCA Pro	GAA Glu 705	GCT Ala	GAT Asp	2349
40	TCT Ser	CTG Leu	TCA Ser 710	Cys	CTG Leu	CAG Gln	GAA Glu	GGA Gly 715	Gln	TGT Cys	GAA Glu	AAT Asn	GAT Asp 720	CCA Pro	AAA Lys	AGC Ser	2397
45	AAA Lys	AAA Lys 725	Val	TCA Ser	GAT Asp	ATA Ile	AAA Lys 730	Glu	GAG Glu	GTC Val	TTG Leu	GCT Ala 735	Ala	GCA Ala	TGT Cys	CAC His	2445
50	CCA Pro 740	Val	CAA Gln	CAC His	TCA Ser	AAA Lys 745	Val	GAA Glu	TAC Tyr	AGT Ser	GAT Asp 750	Thr	GAC Asp	TTT Phe	CAA Gln	TCC Ser 755	2493
FF						Tyr					Ala					TTA Leu	2541
55	ACT Thr	CCI Pro	ACT Thr	TCC Ser 775	Lys	GAT Asp	GTT Val	CTG Leu	TCA Ser 780	Asn	CTA Leu	GTC Val	ATG Met	ATT Ile 785	Ser	AGA Arg	2589
60	GGC Gly	: AAA	A GAA Glu 790	ı Ser	TAC Tyr	Lys	ATC Met	TCA Ser 795	Asp	AAC Lys	CTC Lev	AAA Lys	GGT Gly 800	Asr	AAT Asr	TAT Tyr	2637

5	GAA Glu	TCT Ser 805	GAT Asp	GTT Val	GAA Glu	TTA Leu	ACC Thr 810	AAA Lys	AAT Asn	ATT Ile	CCC Pro	ATG Met 815	GAA Glu	AAG Lys	AAT Asn	CAA Gln	2685
10	GAT Asp 820	GTA Val	TGT Cys	GCT Ala	TTA Leu	AAT Asn 825	GAA Glu	AAT Asn	TAT Tyr	AAA Lys	AAC Asn 830	GTT Val	GAG Glu	CTG Leu	TTG Leu	CCA Pro 835	2733
10	CCT Pro	GAA Glu	AAA Lys	TAC Tyr	ATG Met 840	AGA Arg	GTA Val	GCA Ala	TCA Ser	CCT Pro 845	TCA Ser	AGA Arg	AAG Lys	GTA Val	CAA Gln 850	TTC Phe	2781
15	AAC Asn	CAA Gln	AAC Asn	ACA Thr 855	AAT Asn	CTA Leu	AGA Arg	GTA Val	ATC Ile 860	CAA Gln	AAA Lys	AAT Asn	CAA Gln	GAA Glu 865	GAA Glu	ACT Thr	2829
20	ACT Thr	TCA Ser	ATT Ile 870	TCA Ser	AAA Lys	ATA Ile	ACT Thr	GTC Val 875	AAT Asn	CCA Pro	GAC Asp	TCT Ser	GAA Glu 880	GAA Glu	CTT Leu	TTC Phe	2877
25	TCA Ser	GAC Asp 885	AAT Asn	GAG Glu	AAT Asn	AAT Asn	TTT Phe 890	GTC Val	TTC Phe	CAA Gln	GTA Val	GCT Ala 895	AAT Asn	GAA Glu	AGG Arg	AAT Asn	2925
2.0	AAT Asn 900	CTT Leu	GCT Ala	TTA Leu	GGA Gly	AAT Asn 905	ACT Thr	AAG Lys	GAA Glu	CTT Leu	CAT His 910	GAA Glu	ACA Thr	GAC Asp	TTG Leu	ACT Thr 915	2973
30	TGT Cys	GTA Val	AAC Asn	GAA Glu	CCC Pro 920	ATT Ile	TTC Phe	AAG Lys	AAC Asn	TCT Ser 925	ACC Thr	ATG Met	GTT Val	TTA Leu	TAT Tyr 930	GGA Gly	3021
35	GAC Asp	ACA Thr	GGT Gly	GAT Asp 935	AAA Lys	CAA Gln	GCA Ala	ACC Thr	CAA Gln 940	GTG Val	TCA Ser	ATT Ile	AAA Lys	AAA Lys 945	GAT Asp	TTG Leu	3069
40	GTT Val	TAT Tyr	GTT Val 950	Leu	GCA Ala	GAG Glu	GAG Glu	AAC Asn 955	Lys	AAT Asn	AGT Ser	GTA Val	AAG Lys 960	Gln	CAT His	ATA Ile	3117
45	AAA Lys	ATG Met 965	Thr	CTA Leu	GGT Gly	CAA Gln	GAT Asp 970	Leu	AAA Lys	TCG Ser	GAC Asp	ATC Ile 975	Ser	TTG Leu	AAT Asn	ATA Ile	3165
50	GAT Asp 980	Lys	ATA Ile	CCA Pro	GAA Glu	AAA Lys 985	Asn	AAT Asn	GAT Asp	TAC Tyr	Met	Asp	AAA Lys	TGG Trp	GCA Ala	GGA Gly 995	3213
30	CTC Leu	TTA Leu	GGT Gly	CCA	ATT Ile 1000	Ser	AAT Asn	CAC His	AGT Ser	TTT Phe 1005	Gly	GGT Gly	AGC Ser	TTC Phe	AGA Arg	ACA Thr	3261
55	GCT Ala	TCA Ser	TAA I	AAG Lys 1015	Glu	ATC	: AAG : Lys	CTC Leu	TCT Ser 1020	Glu	CAT His	AAC Asr	ATT	AAG Lys	Lys	AGC Ser	3309
60	AAA Lys	ATC Met	TTC Phe	Phe	AAA Lys	GAT Asp	TATT Ile	GAA Glu 1035	ı Glu	CAA Glm	TAT Tyr	CCI Pro	TACT Thr	Ser	TTA Leu	GCT Ala	3357

F	TGT GTT GAA ATT GTA AAT ACC TTG GCA TTA GAT AAT CAA AAG AAA CTG Cys Val Glu Ile Val Asn Thr Leu Ala Leu Asp Asn Gln Lys Lys Leu 1045 1050 1055	3405
5	AGC AAG CCT CAG TCA ATT AAT ACT GTA TCT GCA CAT TTA CAG AGT AGT Ser Lys Pro Gln Ser Ile Asn Thr Val Ser Ala His Leu Gln Ser Ser 1060 1065 1070 1075	3453
10	GTA GTT GTT TCT GAT TGT AAA AAT AGT CAT ATA ACC CCT CAG ATG TTA Val Val Val Ser Asp Cys Lys Asn Ser His Ile Thr Pro Gln Met Leu 1080 1085 1090	3501
15	TTT TCC AAG CAG GAT TTT AAT TCA AAC CAT AAT TTA ACA CCT AGC CAA Phe Ser Lys Gln Asp Phe Asn Ser Asn His Asn Leu Thr Pro Ser Gln 1095 1100 1105	3549
20	AAG GCA GAA ATT ACA GAA CTT TCT ACT ATA TTA GAA GAA TCA GGA AGT Lys Ala Glu Ile Thr Glu Leu Ser Thr Ile Leu Glu Glu Ser Gly Ser 1110 1115 1120	3597
25	CAG TTT GAA TTT ACT CAG TTT AGA AAG CCA AGC TAC ATA TTG CAG AAG Gln Phe Glu Phe Thr Gln Phe Arg Lys Pro Ser Tyr Ile Leu Gln Lys 1125 1130 1135	3645
23	AGT ACA TTT GAA GTG CCT GAA AAC CAG ATG ACT ATC TTA AAG ACC ACT Ser Thr Phe Glu Val Pro Glu Asn Gln Met Thr Ile Leu Lys Thr Thr 1140 1145 1150 1155	3693
30	TCT GAG GAA TGC AGA GAT GCT GAT CTT CAT GTC ATA ATG AAT GCC CCA Ser Glu Glu Cys Arg Asp Ala Asp Leu His Val Ile Met Asn Ala Pro 1160 1165 1170	3741
35	TCG ATT GGT CAG GTA GAC AGC AGC AAG CAA TTT GAA GGT ACA GTT GAA Ser Ile Gly Gln Val Asp Ser Ser Lys Gln Phe Glu Gly Thr Val Glu 1175 1180 1185	3789
40	ATT AAA CGG AAG TTT GCT GGC CTG TTG AAA AAT GAC TGT AAC AAA AGT Ile Lys Arg Lys Phe Ala Gly Leu Leu Lys Asn Asp Cys Asn Lys Ser 1190 1195 1200	3837
45	GCT TCT GGT TAT TTA ACA GAT GAA AAT GAA GTG GGG TTT AGG GGC TTT Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe Arg Gly Phe 1205 1210 1215	3885
10	TAT TCT GCT CAT GGC ACA AAA CTG AAT GTT TCT ACT GAA GCT CTG CAA Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu Ala Leu Gln 1220 1225 1230 1235	3933
50	AAA GCT GTG AAA CTG TTT AGT GAT ATT GAG AAT ATT AGT GAG GAA ACT Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser Glu Glu Thr 1240 1245 1250	3981
55	TCT GCA GAG GTA CAT CCA ATA AGT TTA TCT TCA AGT AAA TGT CAT GAT Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Lys Cys His Asp 1255 1260 1265	4029
60	TCT GTT GTT TCA ATG TTT AAG ATA GAA AAT CAT AAT GAT AAA ACT GTA Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp Lys Thr Val 1270 1275 1280	4077
	AGT GAA AAA AAT AAA TGC CAA CTG ATA TTA CAA AAT AAT ATT GAA	4125

	Ser Glu 1285	_	Asn	Asn		Cys .290	Gln	Leu	Ile		Gln .295	Asn	Asn	Ile	Glu	
5	ATG ACT Met Thr 1300			Thr					Ile					Lys		4173
10	AAT ACT Asn Thr		Asn					Tyr					Arg			4221
15	CAT AAC His Asn	Leu					Ser					Asn				4269
20	TGT ATT Cys Ile					Thr					Thr					4317
20	ATA TGT Ile Cys 1365	Leu			Ser					Lys						4365
25	ATT AAA Ile Lys 1380			Leu					Phe					Lys		4413
30	CAA GAA Gln Glu		Cys					Ser					Leu			4461
35	ACT AAA Thr Lys	Thr					Lys					Ser				4509
40	TTT CAC		Ala			Lys					Ala					4557
40	AAT AAA Asn Lys 1445	lle		Asn		Phe	Asp	Gln	Lys	Pro	Glu					4605
45	TTT TCC Phe Ser 1460			Ser					Asp					Lys		4653
50	GAC ATT Asp Ile		Ser					Asp					Lys			4701
55	AAA GAA Lys Gli						Thr					Val				4749
60	GGA CAA Gly Glr		Glu			Glu		Ile			Pro					4797
30	TTT CAT															4845

1525 1530 1535

5	GAC AAA GTG AAA AAC CTT TTT GAT GAA AAA GAG CAA GGT ACT AGT GAA Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly Thr Ser Glu 1540 1545 1550 1555	4893
10	ATC ACC AGT TTT AGC CAT CAA TGG GCA AAG ACC CTA AAG TAC AGA GAG  Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys Tyr Arg Glu  1560 1565 1570	4941
1.5	GCC TGT AAA GAC CTT GAA TTA GCA TGT GAG ACC ATT GAG ATC ACA GCT Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu Ile Thr Ala 1575 1580 1585	4989
15	GCC CCA AAG TGT AAA GAA ATG CAG AAT TCT CTC AAT AAT GAT AAA AAC Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn Asp Lys Asn 1590 1595 1600	5037
20	CTT GTT TCT ATT GAG ACT GTG GTG CCA CCT AAG CTC TTA AGT GAT AAT Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu Ser Asp Asn 1605 1610 1615	5085
25	TTA TGT AGA CAA ACT GAA AAT CTC AAA ACA TCA AAA AGT ATC TTT TTG Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser Ile Phe Leu 1620 1625 1630 1635	5133
30	AAA GTT AAA GTA CAT GAA AAT GTA GAA AAA GAA ACA GCA AAA AGT CCT Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala Lys Ser Pro 1640 1645 1650	5181
2.5	GCA ACT TGT TAC ACA AAT CAG TCC CCT TAT TCA GTC ATT GAA AAT TCA Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile Glu Asn Ser 1655 1660 1665	5229
35	GCC TTA GCT TTT TAC ACA AGT TGT AGT AGA AAA ACT TCT GTG AGT CAG Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser Val Ser Gln 1670 1675 1680	5277
40	ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT GAT Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe Asp 1685 1690 1695	5325
45	GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT TTG Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr Leu 1700 1705 1710 1715	5373
50	TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT CAT Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn His 1720 1725 1730	5421
	CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT AAC Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser Met Ser Asn 1735 1740 1745	5469
55	AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT CTC Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr Leu 1750 1755 1760	5517
60	TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT GTT Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn Val 1765 1770 1775	5565

5	GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys 1780 1785 1790 1795	5613
10	GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu 1800 1805 1810	5661
10	GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT Glu Leu Val Thr Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile 1815 1820 1825	5709
15	AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala 1830 1835 1840	5757
20	TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glu Thr Ile 1845 1850 1855	5805
25	AAA AAA GTG AAA GAC ATA TTT ACA GAC AGT TTC AGT AAA GTA ATT AAG Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys 1860 1865 1870 1875	5853
	GAA AAC AAC GAG AAT AAA TCA AAA ATT TGC CAA ACG AAA ATT ATG GCA Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys Ile Met Ala 1880 1885 1890	5901
30	GGT TGT TAC GAG GCA TTG GAT GAT TCA GAG GAT ATT CTT CAT AAC TCT Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu His Asn Ser 1895 1900 1905	5949
35	CTA GAT AAT GAT GAA TGT AGC ACG CAT TCA CAT AAG GTT TTT GCT GAC Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val Phe Ala Asp 1910 1915 1920	5997
40	ATT CAG AGT GAA GAA ATT TTA CAA CAT AAC CAA AAT ATG TCT GGA TTG Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met Ser Gly Leu 1925 1930 1935	6045
45	GAG AAA GTT TCT AAA ATA TCA CCT TGT GAT GTT AGT TTG GAA ACT TCA Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu Glu Thr Ser 1940 1945 1950 1955	6093
5.0	GAT ATA TGT AAA TGT AGT ATA GGG AAG CTT CAT AAG TCA GTC TCA TCT Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser Val Ser Ser 1960 1965 1970	6141
50	GCA AAT ACT TGT GGG ATT TTT AGC ACA GCA AGT GGA AAA TCT GTC CAG Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys Ser Val Gln 1975 1980 1985	6189
55	GTA TCA GAT GCT TCA TTA CAA AAC GCA AGA CAA GTG TTT TCT GAA ATA Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe Ser Glu Ile 1990 1995 2000	6237
60	GAA GAT AGT ACC AAG CAA GTC TTT TCC AAA GTA TTG TTT AAA AGT AAC Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe Lys Ser Asn 2005 2010 2015	6285

5	GAA CAT TCA GAC CAG CTC ACA AGA GAA GAA AAT ACT GCT ATA CGT ACT Glu His Ser Asp Gln Leu Thr Arg Glu Glu Asn Thr Ala Ile Arg Thr 2020 2025 2030 2035	6333
J	CCA GAA CAT TTA ATA TCC CAA AAA GGC TTT TCA TAT AAT GTG GTA AAT Pro Glu His Leu Ile Ser Gln Lys Gly Phe Ser Tyr Asn Val Val Asn 2040 2045 2050	6381
10	TCA TCT GCT TTC TCT GGA TTT AGT ACA GCA AGT GGA AAG CAA GTT TCC Ser Ser Ala Phe Ser Gly Phe Ser Thr Ala Ser Gly Lys Gln Val Ser 2055 2060 2065	6429
15	ATT TTA GAA AGT TCC TTA CAC AAA GTT AAG GGA GTG TTA GAG GAA TTT  Ile Leu Glu Ser Ser Leu His Lys Val Lys Gly Val Leu Glu Glu Phe  2070 2075 2080	6477
20	GAT TTA ATC AGA ACT GAG CAT AGT CTT CAC TAT TCA CCT ACG TCT AGA Asp Leu Ile Arg Thr Glu His Ser Leu His Tyr Ser Pro Thr Ser Arg 2085 2090 2095	6525
25	CAA AAT GTA TCA AAA ATA CTT CCT CGT GTT GAT AAG AGA AAC CCA GAG Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg Asn Pro Glu 2100 2105 2110 2115	6573
	CAC TGT GTA AAC TCA GAA ATG GAA AAA ACC TGC AGT AAA GAA TTT AAA His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys Glu Phe Lys 2120 2125 2130	6621
30	TTA TCA AAT AAC TTA AAT GTT GAA GGT GGT TCT TCA GAA AAT AAT CAC Leu Ser Asn Asn Leu Asn Val Glu Gly Gly Ser Ser Glu Asn Asn His 2135 2140 2145	6669
35	TCT ATT AAA GTT TCT CCA TAT CTC TCT CAA TTT CAA CAA GAC AAA CAA Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln Asp Lys Gln 2150 2155 2160	6717
40	CAG TTG GTA TTA GGA ACC AAA GTC TCA CTT GTT GAG AAC ATT CAT GTT Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn Ile His Val 2165 2170 2175	6765
45	TTG GGA AAA GAA CAG GCT TCA CCT AAA AAC GTA AAA ATG GAA ATT GGT Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met Glu Ile Gly 2180 2185 2190 2195	6813
	AAA ACT GAA ACT TTT TCT GAT GTT CCT GTG AAA ACA AAT ATA GAA GTT Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn Ile Glu Val 2200 2205 2210	6861
50	TGT TCT ACT TAC TCC AAA GAT TCA GAA AAC TAC TTT GAA ACA GAA GCA Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu Thr Glu Ala 2215 2220 2225	6909
55	GTA GAA ATT GCT AAA GCT TTT ATG GAA GAT GAT GAA CTG ACA GAT TCT Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu Thr Asp Ser 2230 2235 2240	6957
60	AAA CTG CCA AGT CAT GCC ACA CAT TCT CTT TTT ACA TGT CCC GAA AAT Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys Pro Glu Asn 2245 2250 2255	7005
	GAG GAA ATG GTT TTG TCA AAT TCA AGA ATT GGA AAA AGA AGA GGA GAG	7053

	Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg Arg Gly Glu 2260 2265 2270 2275	
5	CCC CTT ATC TTA GTG GGA GAA CCC TCA ATC AAA AGA AAC TTA TTA AAT Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn Leu Leu Asn 2280 2285 2290	7101
10	GAA TTT GAC AGG ATA ATA GAA AAT CAA GAA AAA TCC TTA AAG GCT TCA Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu Lys Ala Ser 2295 2300 2305	7149
15	AAA AGC ACT CCA GAT GGC ACA ATA AAA GAT CGA AGA TTG TTT ATG CAT Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu Phe Met His 2310 2315 2320	7197
2.0	CAT GTT TCT TTA GAG CCG ATT ACC TGT GTA CCC TTT CGC ACA ACT AAG His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg Thr Thr Lys 2325 2330 2335	7245
20	GAA CGT CAA GAG ATA CAG AAT CCA AAT TTT ACC GCA CCT GGT CAA GAA Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro Gly Gln Glu 2340 2345 2350 2355	7293
25	TTT CTG TCT AAA TCT CAT TTG TAT GAA CAT CTG ACT TTG GAA AAA TCT Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu Glu Lys Ser 2360 2365 2370	7341
30	TCA AGC AAT TTA GCA GTT TCA GGA CAT CCA TTT TAT CAA GTT TCT GCT Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln Val Ser Ala 2375 2380 2385	7389
35	ACA AGA AAT GAA AAA ATG AGA CAC TTG ATT ACT ACA GGC AGA CCA ACC Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly Arg Pro Thr 2390 2395 2400	7437
4.0	AAA GTC TTT GTT CCA CCT TTT AAA ACT AAA TCG CAT TTT CAC AGA GTT Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe His Arg Val 2405 2410 2415	7485
40	GAA CAG TGT GTT AGG AAT ATT AAC TTG GAG GAA AAC AGA CAA AAG CAA Glu Gln Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg Gln Lys Gln 2420 2425 2430 2435	7533
45	AAC ATT GAT GGA CAT GGC TCT GAT GAT AGT AAA AAT AAG ATT AAT GAC Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys Ile Asn Asp 2440 2445 2450	7581
50	AAT GAG ATT CAT CAG TTT AAC AAA AAC AAC TCC AAT CAA GCA GCT Asn Glu Ile His Gln Phe Asn Lys Asn Asn Ser Asn Gln Ala Ala 2455 2460 2465	7629
55	GTA ACT TTC ACA AAG TGT GAA GAA GAA CCT TTA GAT TTA ATT ACA AGT Val Thr Phe Thr Lys Cys Glu Glu Pro Leu Asp Leu Ile Thr Ser 2470 2475 2480	7677
60	CTT CAG AAT GCC AGA GAT ATA CAG GAT ATG CGA ATT AAG AAG AAA CAA Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys Lys Lys Gln 2485 2490 2495	7725
60	AGG CAA CGC GTC TTT CCA CAG CCA GGC AGT CTG TAT CTT GCA AAA ACA Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu Ala Lys Thr	7773

to the free state

TTA AAG AAT GGC AGA CTG ACA GTT GGT CAG AAG ATT ATT CTT CAT GGA

Leu Lys Asn Gly Arg Leu Thr Val Gly Gln Lys Ile Ile Leu His Gly

5				Val					Ala	TGT Cys 2765				Glu	_		8541
10			Leu					Ser		AAC Asn			Arg				8589
10		Tyr					Phe			GAC Asp		Arg					8637
15	Pro					Phe				GGA Gly	Asn						8685
20					Arg					CAG Gln					Thr		8733
25				Tyr					Glu	AGA Arg 2845				Lys			8781
2.0			Tyr					Gln		AGA Arg			Ala				8829
30		Ile					Glu			GAA Glu		Asn					8877
35	Tyr					Ala				CAG Gln	Gln						8925
40		Gly			Leu					AAG Lys					Pro		8973
45	TAC Tyr	CTT Leu	GAG Glu	Gly	TAT Tyr 2920	TTC Phe	AGT Ser	GAA Glu	Glu	CAG Gln 2925	TTA Leu	AGA Arg	GCC Ala	Leu	AAT Asn 2930	AAT Asn	9021
50			Gln					Lys					Ile		Leu	GAA Glu	9069
30		Arg		Ala			Ser					Glu				TCA Ser	9117
55	Arg					Val				CGT Arg	Ile					AAA Lys	9165
60		Glu			Ser					Ile		Arg				GAT Asp 2995	9213

5	TTA TAT TCT CTG TTA ACA GAA GGA AAG AGA TAC AGA ATT TAT CAT CTT Leu Tyr Ser Leu Leu Thr Glu Gly Lys Arg Tyr Arg Ile Tyr His Leu 3000 3005 3010	9261
5	GCA ACT TCA AAA TCT AAA AGT AAA TCT GAA AGA GCT AAC ATA CAG TTA Ala Thr Ser Lys Ser Lys Ser Glu Arg Ala Asn Ile Gln Leu 3015 3020 3025	9309
10	GCA GCG ACA AAA AAA ACT CAG TAT CAA CAA CTA CCG GTT TCA GAT GAA Ala Ala Thr Lys Lys Thr Gln Tyr Gln Gln Leu Pro Val Ser Asp Glu 3030 3035 3040	9357
15	ATT TTA TTT CAG ATT TAC CAG CCA CGG GAG CCC CTT CAC TTC AGC AAA Ile Leu Phe Gln Ile Tyr Gln Pro Arg Glu Pro Leu His Phe Ser Lys 3045 3050 3055	9405
20	TTT TTA GAT CCA GAC TTT CAG CCA TCT TGT TCT GAG GTG GAC CTA ATA  Phe Leu Asp Pro Asp Phe Gln Pro Ser Cys Ser Glu Val Asp Leu Ile  3060 3065 3070 3075	9453
25	GGA TTT GTC GTT TCT GTT GTG AAA AAA ACA GGA CTT GCC CCT TTC GTC Gly Phe Val Val Ser Val Val Lys Lys Thr Gly Leu Ala Pro Phe Val 3080 3085 3090	9501
25	TAT TTG TCA GAC GAA TGT TAC AAT TTA CTG GCA ATA AAG TTT TGG ATA Tyr Leu Ser Asp Glu Cys Tyr Asn Leu Leu Ala Ile Lys Phe Trp Ile 3095 3100 3105	9549
30	GAC CTT AAT GAG GAC ATT ATT AAG CCT CAT ATG TTA ATT GCT GCA AGC Asp Leu Asn Glu Asp Ile Ile Lys Pro His Met Leu Ile Ala Ala Ser 3110 3120	9597
35	AAC CTC CAG TGG CGA CCA GAA TCC AAA TCA GGC CTT CTT ACT TTA TTT Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu Thr Leu Phe 3125 3130 3135	9645
40	GCT GGA GAT TTT TCT GTG TTT TCT GCT AGT CCA AAA GAG GGC CAC TTT Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu Gly His Phe 3140 3145 3150 3155	9693
45	CAA GAG ACA TTC AAC AAA ATG AAA AAT ACT GTT GAG AAT ATT GAC ATA Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn Ile Asp Ile 3160 3165 3170	9741
10	CTT TGC AAT GAA GCA GAA AAC AAG CTT ATG CAT ATA CTG CAT GCA AAT Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu His Ala Asn 3175 3180 3185	9789
50	GAT CCC AAG TGG TCC ACC CCA ACT AAA GAC TGT ACT TCA GGG CCG TAC Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser Gly Pro Tyr 3190 3195 3200	9837
55	ACT GCT CAA ATC ATT CCT GGT ACA GGA AAC AAG CTT CTG ATG TCT TCT Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu Met Ser Ser 3205 3210 3215	9885
60	CCT AAT TGT GAG ATA TAT TAT CAA AGT CCT TTA TCA CTT TGT ATG GCC Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu Cys Met Ala 3220 3225 3230 3235	9933
	AAA AGG AAG TCT GTT TCC ACA CCT GTC TCA GCC CAG ATG ACT TCA AAG	9981

	Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met Thr Ser Lys 3240 3245 3250	
5	TCT TGT AAA GGG GAG AAA GAG ATT GAT GAC CAA AAG AAC TGC AAA AAG Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn Cys Lys 3255 3260 3265	10029
10	AGA AGA GCC TTG GAT TTC TTG AGT AGA CTG CCT TTA CCT CCA CCT GTT Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro Pro Pro Val 3270 3280	10077
15	AGT CCC ATT TGT ACA TTT GTT TCT CCG GCT GCA CAG AAG GCA TTT CAG Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys Ala Phe Gln 3285 3290 3295	10125
20	CCA CCA AGG AGT TGT GGC ACC AAA TAC GAA ACA CCC ATA AAG AAA AAA Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile Lys Lys 3300 3315	10173
20	GAA CTG AAT TCT CCT CAG ATG ACT CCA TTT AAA AAA TTC AAT GAA ATT Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe Asn Glu Ile 3320 3330	10221
25	TCT CTT TTG GAA AGT AAT TCA ATA GCT GAC GAA GAA CTT GCA TTG ATA Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu Ala Leu Ile 3335 3340 3345	10269
30	AAT ACC CAA GCT CTT TTG TCT GGT TCA ACA GGA GAA AAA CAA TTT ATA Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys Gln Phe Ile 3350 3355 3360	10317
35	TCT GTC AGT GAA TCC ACT AGG ACT GCT CCC ACC AGT TCA GAA GAT TAT Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser Glu Asp Tyr 3365 3370 3375	10365
40	CTC AGA CTG AAA CGA CGT TGT ACT ACA TCT CTG ATC AAA GAA CAG GAG Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys Glu Gln Glu 3380 3385 3390 3395	10413
40	AGT TCC CAG GCC AGT ACG GAA GAA TGT GAG AAA AAT AAG CAG GAC ACA Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys Gln Asp Thr 3400 3405 3410	10461
45	ATT ACA ACT AAA AAA TAT ATC TAA Ile Thr Thr Lys Lys Tyr Ile 3415	10485
50	(2) INFORMATION FOR SEQ ID NO:9:	
55	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 3418 amino acids</li> <li>(B) TYPE: amino acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
60	<ul><li>(ii) MOLECULE TYPE: protein</li><li>(v) FRAGMENT TYPE: internal</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:</li></ul>	

	Met 1	Pro	Ile	Gly	Ser 5	Lys	Glu	Arg	Pro	Thr 10	Phe	Phe	Glu	Ile	Phe 15	Lys
5		Arg	Cys	Asn 20	Lys	Ala	Asp	Leu	Gly 25	Pro	Ile	Ser	Leu	Asn 30	Trp	Phe
J	Glu	Glu	Leu 35		Ser	Glu	Ala	Pro 40		Tyr	Asn	Ser	Glu 45	Pro	Ala	Glu
	Glu	Ser 50		His	Lys	Asn	Asn 55		Tyr	Glu	Pro	Asn 60	Leu	Phe	Lys	Thr
10	Pro 65	Gln	Arg	Lys	Pro	Ser 70		Asn	Gln	Leu	Ala 75	Ser	Thr	Pro	Ile	Ile 80
		Lys	Glu	Gln	Gly 85	Leu	Thr	Leu	Pro	Leu 90	Tyr	Gln	Ser	Pro	Val 95	Lys
15	Glu	Leu	Asp	Lys 100	Phe	Lys	Leu	Asp	Leu 105	Gly	Arg	Asn	Val	Pro 110	Asn	Ser
	_	His	115					120					125			
		Val 130					135					140				
20	145	Leu				150					155					160
		Gly			165					170					175	
25		Lys		180					185					190		
		Trp	195					200					205			
		Ile 210		_			215					220				
30	225	Thr				230					235					240
	-	Lys			245					250					255	
35		Gln		260					265					270		
		Phe	275					280					285			
4.0		Val 290 Asp					295					300				
40	305					310					315					320
		. Lys Ala			325					330					335	
45		Phe		340					345					350		
		Val	355					360					365			
50		370 Glu					375					380				
30	385					390					395					400
		Ser			405					410					415	
55		ı Lys		420					425					430		
		e Ser	435	i				440					445			
60		450 Asn	ı				455	;				460	)			
	465 Il∈		Ala	. Val	Lys	470 Glr		ı Ile	Ser	Gly	475 Thr		Pro	val	Ala	480 Ser

	Ser	Dhe	Gln	Glv	485 Tle	Lvs	Lvs	Ser	Ile	490 Phe	Ara	Ile	Arq	Glu	495 Ser	Pro
_				500					505					510		
5	_		Thr 515					520					525			
		530	Lys				535					540				
10	Val 545	Cys	Ser	Gln	Lys	Glu 550	Asp	Ser	Leu	Cys	Pro 555	Asn	Leu	Ile	Asp	Asn 560
_ •		Ser	Trp	Pro	Ala 565	Thr	Thr	Thr	Gln	Asn 570	Ser	Val	Ala	Leu	Lys 575	Asn
			Leu	580					585					590		
15			His 595					600					605			
		610	Ser				615					620				
20	Phe 625	Glu	Ala	Pro	Leu	Thr 630	Phe	Ala	Asn	Ala	Asp 635	Ser	Gly	Leu	Leu	His 640
20	Ser		Val		645					650					655	
	Leu	Ser	Leu	Thr 660	Ser	Ser	Phe	Gly	Thr 665	Ile	Leu	Arg	Lys	Cys 670	Ser	Arg
25	Asn	Glu	Thr 675		Ser	Asn	Asn	Thr 680	Val	Ile	Ser	Gln	Asp 685	Leu	Asp	Tyr
	-	690	Ala	_	_		695					700				
30	Glu 705	Ala	Asp	Ser	Leu	Ser 710	Cys	Leu	Gln	Glu	Gly 715	Gln	Cys	Glu	Asn	Asp 720
		-	Ser		725					730					735	
			His	740					745					750		
35			Ser 755					760					765			
		770					775					780				
40	785					790					795					Gly 800
			Tyr		805					810					815	
			Gln	820					825					830		
45			835					840					845			Lys
		850					855					860				Gln
50	Glu 865		Thr	Thr	Ser	Ile 870	Ser	Lys	Ile	Thr	Val 875		Pro	Asp	Ser	Glu 880
	Glu	Leu			885					890					895	
				900					905					910		Thr
55			915					920					925			Val
		930	)				935					940				Lys
60	Lys 945		Leu	Val	Tyr	Val 950		Ala	Glu	Glu	Asn 955		Asn	Ser	· Val	Lys 960
	Glr	n His	: Ile	. Lys	Met 965		Leu	Gly	Gln	970		Lys	Ser	Asp	975	Ser

	Leu	Asn	Ile	Asp 980	Lys	Ile	Pro	Glu	Lys 985	Asn	Asn	Asp	Tyr	Met 990	Asp	Lys
5	Trp	Ala	Gly 995	Leu	Leu	Gly	Pro	Ile 1000		Asn	His	Ser	Phe 1005		Gly	Ser
	Phe	Arg 1010		Ala	Ser	Asn	Lys 1015	Glu	Ile	Lys	Leu	Ser 1020		His	Asn	Ile
	Lys 1025	-	Ser	Lys	Met	Phe 1030		Lys	Asp	Ile	Glu 1035		Gln	Tyr	Pro	Thr 104
10				-	1045	i		Val		1050	)				1055	5
				1060	)			Ser	1065	5				1070	)	
15			1075	5				Asp 1080	1				1085	5		
		1090	)				1095					1100	)			
	1105	5				1110	)	Thr			1115	5				112
20		_			1125	5		Thr		1130	)				1135	5
			_	1140	)			Val	1145	5				1150	)	
25			1155	5				Arg 1160	)				1165	5		
		1170	)				1175					1180	)			
	Thr 1185		Glu	Ile	Lys	Arg 1190	_	Phe	Ala	Gly	Leu 119		Lys	Asn	Asp	Cys 120
30		-			1205	5		Leu		1210	)				1215	5
	_	_		1220	)			Gly	1225	5				1230	)	
35			1235	5				Leu 1240	)				124	5		
		1250	3				125					1260	)			
	126	5	-			1270	)	Met			127	5				128
40	-				128	5		Asn		129	0				129	5
				1300	0			Thr	130	5				131	0	
45			131	5				Glu 1320	)				132	5		
		133	0				133					134	0			
F.0	134	5				135	С	Asp			135	5				136
50					136	5		Leu		137	0				137	5
				138	0			Leu	138	5				139	0	
55		-	139	5			_	1400	)				140	5		Gln
		141	0		_		141					142	0			
<b>C</b> O	142	5				143	0	Ser			143	5				144
60					144	5		Asn		145	0				145	5
	Leu	His	Asn	Phe	ser	ьeu	Asn	ser	Glu	ьeu	HIS	ser	Asp	тте	arg	Lys

				1460					1465	·				1470		
		_	1475					1480					1485	;		
5	_	1490	)	Lys			1495					1500				
	Thr 1505		Gln	Gly	Gln	Pro 1510		Arg	Asp	Glu	Lys 1515		Lys	Glu	Pro	Thr 152
10			_	Phe	1525	,			_	1530					1535	5
				Asp 1540	)				1545	5				1550	ı	
			1555					1560	)				1565	5		
15	-	1570	)	Ala			1575	i				1580	)			
	1585	5		Ala		1590	)				1595					160
20				Leu	1605	5				1610	1				1615	5
		_		Leu 1620	)	_			1625	5				1630	)	
			1635			_		1640	)				1645	5		
25	_	1650	)	Ala		_	1655	; ;				1660	)			
	1665	5		Ala		1670	)				1675	;				168
30				Thr	1685	5				1690	)				1695	5
			-	Gly 1700	)			_	1705	5				1710	)	
2.5		-	1715					1720	)				1725	5		
35	_	1730	)	Leu			1735	5				1740	)			
	174	5		Ser	_	1750	) _				1755	5				176
40	-	-		Ser	1765	5	_		_	1770	)				1775	5
	_			Glu 1780	)				1785	5				1790	)	
45			1799	Asp Glu				1800	)				1805	5		
40	-	1810	)	Lys			1815	5				1820	)			
	182	5		Phe		1830	)				1835	5				184
50				Lys	1845	5				1850	)				185	5
				1860 Glu	) _		-	_	1869	5		_		1870	)	
55			1879					1880	)				1885	5		
		189	0	Leu			1899	5				1900	)			
	190	5		Ile		1910	0				1915	5				192
60			_	Glu	192	5				1930	)				193	5
	J	J-1		1940	-	. •		-1-	194			-1~	F	1950		

	Glu	Thr	Ser 1955		Ile	Cys	Lys	Cys 1960		Ile	Gly	Lys	Leu 1965		Lys	Ser
5	Val	Ser 1970	Ser		Asn	Thr	Cys 1975	Gly		Phe	Ser	Thr	Ala		Gly	Lys
J	Ser 1985	Val		Val		Asp 1990	Ala	Ser	Leu	Gln	Asn 1995		Arg	Gln	Val	Phe 200
					2005	5		Lys		2010	)				2015	i
10	_			2020	)			Gln	2025	5				2030	)	
		_	2035	5				Ile 2040	1				2045	5		
15		2050	)				2055	Ser Ser				2060	)			
	206	5				2070	)				2075	5				208
					2085	5		Thr		2090	С				2095	5
20				2100	)			Lys Ser	2109	5				2110	)	
			2119	5				2120	)				2125	5		
25	Glu	Phe 213	-	Leu	Ser	Asn	Asn 213	Leu 5	Asn	Val	Glu	Gly 214		Ser	Ser	Glu
	214	5				2150	)	Ser			215	5				216
	Asp	Lys	Gln	Gln	Leu 216		Leu	Gly	Thr	Lys 217		Ser	Leu	Val	Glu 2175	
30				2180	)			Gln	218	5				219	O	
	Glu	Ile	Gly 219		Thr	Glu	Thr	Phe 2200		Asp	Val	Pro	Val 220		Thr	Asn
35		221	0				221					222	0			
	222	5				2230	0	Lys			223	5				224
		_		_	224	5		His		225	0				225	5
40				226	0			Leu	226	5				227	0	
	_	_	227	5				Val 228	)				228	5		
45		229	0				229					230	0			
	Lys 230		Ser	Lys	Ser	Thr 231		Asp	Gly	Thr	Ile 231		Asp	Arg	Arg	Leu 232
			His	His	Val 232	Ser		Glu	Pro	Ile 233	Thr		Val	Pro	Phe 233	
50	Thr	Thr	Lys	Glu 234	_	Gln	Glu	Ile	Gln 234		Pro	Asn	Phe	Thr 235		Pro
	Gly	Gln	Glu 235	Phe		Ser	Lys	Ser 236	His		Tyr	Glu	His 236	Leu		Leu
55	Glu	Lys 237		Ser	Ser	Asn	Leu 237	Ala 5	Val	Ser	Gly	His 238		Phe	Tyr	Gln
33	Val 238	Ser		Thr	Arg	Asn 239	Glu	Lys	Met	Arg	His 239	Leu		Thr	Thr	Gly 240
			Thr	Lys	Val 240		Val	Pro	Pro	Phe 241		Thr	Lys	Ser	His 241	
60	His	Arg	Val	Glu 242	Gln		Val	Arg	Asn 242	Ile		Leu	Glu	Glu 243	Asn	
	Gln	Lys	Gln			Asp	Gly	His			Asp	Asp	Ser			Lys

		2435				2440					2445			
	Ile Asn 245	Asp As	sn Glu			Gln		Asn		Asn 2460		Ser	Asn	Gln
5	Ala Ala 2465	Ala Va	al Thr	Phe '		Lys	Cys		Glu 2475		Pro	Leu	Asp	Leu 248
	Ile Thr	Ser Le	eu Gln 2485		Ala	Arg	Asp	Ile 2490		Asp	Met	Arg	Ile 2495	
10	Lys Lys		g Gln	Arg	Val		Pro 2505		Pro	Gly	Ser	Leu 2510		Leu
	Ala Lys	Thr Se 2515	er Thr	Leu		Arg 2520		Ser	Leu	Lys	Ala 2525		Val	Gly
	Gly Gln 253		co Ser		Cys 2535		His	Lys		Leu 2540		Thr	Tyr	Gly
15	Val Ser 2545	Lys H	is Cys	Ile 2550		Ile	Asn	Ser	Lys 2555		Ala	Glu	Ser	Phe 256
	Gln Phe		2565	5				2570	)				2575	5
20	Lys Gly		ln Leu 580	Ala .	Asp		Gly 2585		Leu	Ile	Pro	Ser 2590		Asp
	Gly Lys	2595				2600					2605	;		
	Gly Val 261	0	_		2615		_			2620	)			
25	Arg Trp 2625			2630					2635					264
	Glu Phe		2645	5				2650	)				2655	5
30	Lys Tyr	2	560				2665	5				2670	)	
	Lys Lys	2675				2680					2685	5		
	Cys Val 269	0			2695					2700	)			
35	Ser Asn 2705	_		2710					2715	;				272
	Leu Thr		272	5				2730	)				273	5
40	Leu Ala	2	740				2745	5				2750	)	
	Leu His	2755				2760	)				2765	5		
4.5	Glu Ala 277	0			2775	5				2780	)			
45	Pro Ala 2785	_		2790					2795	5				280
	Phe Pro		280	5				2810	)				281	5
50	Cys Val	2	820				2825	5				283	0	
	Lys Glu	2835	_		_	2840	)				284	5		
55	285 Leu Phe	0	_	_	2855	5				286	0			
JJ	2865 Thr Lys			2870	)				2875	5				288
	Ala Leu		288	5				2890	)				289	5
60	Asp Pro	2	900				290	5				291	0	
	wah erc	2915	ır nea	Jiu	Cry	2920		501	<u> </u>	<u> Jau</u>	292		9	

	Leu	Asn 2930		His	Arg	Gln	Met 2935		Asn	Asp	Lys	Lys 2940	Gln	Ala	Gln	Ile
	Gln	T.911	Glu	Tle	Δra	Lvs			Glu	Ser	Ala	Glu	Gln	Lvs	Glu	Gln
5	2945		Ciu	110	712 9	2950					2955			-		296
J	234.	, T.211	Ser	Δκα	Δen			Thr	Val	Trp			Ara	Ile	Val	Ser
	Gry	шец	DCI	Arg	2965		1111	1	· u =	2970			5		2975	
	TT:	Cox	Lys	Tara			Λαn	Ser	Val			Ser	Tle	Trp		
	Tyr	ser	гуя			цур	Asp	261	2985		пси	DCI	110	2990	1	
10	A	0	Asp	2980		Cor	T OU	Len			Glv	Taye	Δra			Tle
10	ser	ser			TAT	ser		3000		GIU	O <sub>T</sub> y	טעם	3005		9	
	_		2995 Leu		ml	0				cor	Tage	Car			Δla	Δen
	Tyr			Ala	THE	ser			гуу	per	цуь	3020	J	ria	AIG	ADII
	_	3010	) _			em)	3015		m1	<b>~1</b>				T 011	Dro	17.7
			Leu	Ala	Ala			ьуs	1111	GIII	TAT	- 6111	GIII	ьeu	PIO	304
15	3029	5			_	3030		<b>-</b> 3 -	m	<b>a</b> 1	3035		G3.,	Dxo	T 011	
	Ser	Asp	Glu	Ile			GIn	Пе	Tyr			Arg	GIU	Pro		
				_	3045		_	_	·1	3050		0	G	0	3055	
	Phe	Ser	Lys			Asp	Pro	Asp			Pro	ser	Cys			vai
				3060					3065					3070		
20	Asp	Leu	Ile	Gly	Phe	Val	Val			Val	Lys	Lys			ьeu	АТА
			3075	5				3080					3085			_
	Pro	Phe	Val	Tyr	Leu	Ser	Asp	Glu	Cys	Tyr	Asn			Ala	Ile	Lys
		309					3099					310				
	Phe	Trp	Ile	Asp	Leu	Asn	Glu	Asp	Ile	Ile	Lys	Pro	His	Met	Leu	Ile
25	310	5				311					311					312
	Ala	Ala	Ser	Asn	Leu	Gln	Trp	Arg	Pro	Glu	Ser	Lys	Ser	Gly	Leu	Leu
					312	5				313	0				313	5
	Thr	Leu	Phe	Ala	Gly	Asp	Phe	Ser	Val	Phe	Ser	Ala	Ser	Pro	Lys	Glu
				314	0				314	5				315	0	
30	Glv	His	Phe	Gln	Glu	Thr	Phe	Asn	Lys	Met	Lys	Asn	Thr	Val	Glu	Asn
	1		315					316					316			
	Ile	Asp	Ile	Leu	Cys	Asn	Glu	Ala	Glu	Asn	Lys	Leu	Met	His	Ile	Leu
		317			1		317					318				
	His	Ala	Asn	Asp	Pro	Lys	Trp	Ser	Thr	Pro	Thr	Lys	Asp	Cys	Thr	Ser
35	318					319					319					320
00	Glv	Pro	Tyr	Thr	Ala	Gln	Ile	Ile	Pro	Gly	Thr	Gly	Asn	Lys	Leu	Leu
			-1-		320					321		_			321	
	Met	Ser	Ser	Pro			Glu	Ile	Tyr	Tyr	Gln	Ser	Pro	Leu	Ser	Leu
	1100	501	DCI	322		0,70			322					323		
40	Cvs	Met	Ala	Lvs	Ara	Lvs	Ser	Val			Pro	Val	Ser	Ala	Gln	Met
10	Cyb	1100	323			-7-		324					324			
	Thr	Ser	Lys		Cvs	Lvs	Glv			Glu	Ile	Asp	Asp	Gln	Lys	Asn
	1111	325	0	501	0,0	-1-	325	5	-2-			326	0		-	
	Circ	Lyc	Lys	λκα	Ara	Δla	Leu	Asn	Phe	Len	Ser	Ara	Leu	Pro	Leu	Pro
45	326		цуз	Arg	AT 9	327		1100	1 110		327					328
43			Val	Car	Dro			Thr	Phe	Val			Ala	Ala	Gln	
	PIC	PIO	val	Ser	328		Суз	1111	1110	329					329	
	- דית	Dho	Gln	Dro			Car		Glv			TVY	Glu	Thr		
	Ala	Pne	GIII			ALG	261	Cys	330		цуз	- y -	Q L u	331		
EΛ	<b>.</b>	T	Lys	330		Nan	Cor	Dro			Thr	Dro	Dhe			Phe
50	гÀг	гуѕ			. Leu	ASI	Ser	332		Mec	. 1111	ric	332			1 110
	3	<b>a</b> 1	331 Ile				<i>α</i> 1			Car	· T]	_ ה			Glu	T.e.u
	Asr			ser	ьeu	. ьеи			ASII	361	110	334		Olu	. Olu	Вси
		333	.0	7	ml		333		T 0.11	Cox				Glv	Glu	T.370
			ılle	Asn	inr			пеи	ьeu	. ser			. 1111	Gly	Ģīu	336
55	334	:5				335			mb	. 7	335		D~o	The	Car	
	GIr	Phe	: Ile	ser			Glu	Ser	THE			Alc	PIC	, 1111	337	
	_		_	_	336		_	_	_	337		. ml	- 0			
	Gli	ı Asp	Tyr			Leu	г гув	Arg			ınr	Thr	. ser			ь
				338				_	338		. ~ .			339		T
60	Glı	ı Glr	ı Glu		Ser	Glr	ı Ala			GIU	ı GIV	суя			AST	ггу
			339				_	340		<b>-</b> -			340	5		
	Glr	ı Asp	Thr	: Ile	: Thr	Thr	: Lys	: Lys	Tyr	. 11e	•					

3410 3415

	<b>-</b>	(2) INFORMATION FOR SEQ ID NO:10:	
	5 10	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 10485 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: double</li><li>(D) TOPOLOGY: linear</li></ul>	
		(ii) MOLECULE TYPE: cDNA (ix) FEATURE:	
	15	<ul><li>(A) NAME/KEY: Coding Sequence</li><li>(B) LOCATION: 22910482</li><li>(D) OTHER INFORMATION: BRCA2 (OMI4)</li></ul>	
	20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:	
Train train	25	ACAGATTTGT GACCGGCGCG GTTTTTGTCA GCTTACTCCG GCCAAAAAAG AACTGCACCT	60 120 180 37
	30	GGA TCC AAA GAG AGG CCA ACA TTT TTT GAA ATT TTT AAG ACA CGC TGC Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys Thr Arg Cys 5 10 15	85
: :	35	AAC AAA GCA GAT TTA GGA CCA ATA AGT CTT AAT TGG TTT GAA GAA CTT Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe Glu Glu Leu 20 25 30 35	33
	33	TCT TCA GAA GCT CCA CCC TAT AAT TCT GAA CCT GCA GAA GAA TCT GAA Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu Glu Ser Glu 40 45 50	81
	40	CAT AAA AAC AAC AAT TAC GAA CCA AAC CTA TTT AAA ACT CCA CAA AGG His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr Pro Gln Arg 55 60 65	29
	45	AAA CCA TCT TAT AAT CAG CTG GCT TCA ACT CCA ATA ATA TTC AAA GAG Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile Phe Lys Glu 70 75 80	77
	50	CAA GGG CTG ACT CTG CCG CTG TAC CAA TCT CCT GTA AAA GAA TTA GAT  Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys Glu Leu Asp  85  90  95	25
	<b>-</b> -	AAA TTC AAA TTA GAC TTA GGA AGG AAT GTT CCC AAT AGT AGA CAT AAA  Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser Arg His Lys  100 115	73
	55	AGT CTT CGC ACA GTG AAA ACT AAA ATG GAT CAA GCA GAT GAT GTT TCC Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp Asp Val Ser 120 125 130	21
	60	TGT CCA CTT CTA AAT TCT TGT CTT AGT GAA AGT CCT GTT GTT CTA CAA  Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val Val Leu Gln  135  140  145	69

5	TGT Cys	ACA Thr	CAT His 150	GTA Val	ACA Thr	CCA Pro	CAA Gln	AGA Arg 155	GAT Asp	AAG Lys	TCA Ser	GTG Val	GTA Val 160	TGT Cys	GGG Gly	AGT Ser	717
10	TTG Leu	TTT Phe 165	CAT His	ACA Thr	CCA Pro	AAG Lys	TTT Phe 170	GTG Val	AAG Lys	GGT Gly	CGT Arg	CAG Gln 175	ACA Thr	CCA Pro	AAA Lys	CAT His	765
10	ATT Ile 180	TCT Ser	GAA Glu	AGT Ser	CTA Leu	GGA Gly 185	GCT Ala	GAG Glu	GTG Val	GAT Asp	CCT Pro 190	GAT Asp	ATG Met	TCT Ser	TGG Trp	TCA Ser 195	813
15	AGT Ser	TCT Ser	TTA Leu	GCT Ala	ACA Thr 200	CCA Pro	CCC Pro	ACC Thr	CTT Leu	AGT Ser 205	TCT Ser	ACT Thr	GTG Val	CTC Leu	ATA Ile 210	GTC Val	861
20	AGA Arg	AAT Asn	GAA Glu	GAA Glu 215	GCA Ala	TCT Ser	GAA Glu	ACT Thr	GTA Val 220	TTT Phe	CCT Pro	CAT His	GAT Asp	ACT Thr 225	ACT Thr	GCT Ala	909
25	AAT Asn	GTG Val	AAA Lys 230	AGC Ser	TAT Tyr	TTT Phe	TCC Ser	AAT Asn 235	CAT His	GAT Asp	GAA Glu	AGT Ser	CTG Leu 240	AAG Lys	AAA Lys	AAT Asn	957
2.0	GAT Asp	AGA Arg 245	TTT Phe	ATC Ile	GCT Ala	TCT Ser	GTG Val 250	ACA Thr	GAC Asp	AGT Ser	GAA Glu	AAC Asn 255	ACA Thr	AAT Asn	CAA Gln	AGA Arg	1005
30	GAA Glu 260	GCT Ala	GCA Ala	AGT Ser	CAT His	GGA Gly 265	TTT Phe	GGA Gly	AAA Lys	ACA Thr	TCA Ser 270	GGG Gly	AAT Asn	TCA Ser	TTT Phe	AAA Lys 275	1053
35	GTA Val	AAT Asn	AGC Ser	TGC Cys	AAA Lys 280	GAC Asp	CAC His	ATT Ile	GGA Gly	AAG Lys 285	TCA Ser	ATG Met	CCA Pro	AAT Asn	GTC Val 290	CTA Leu	1101
40	GAA Glu	GAT Asp	GAA Glu	GTA Val 295	TAT Tyr	GAA Glu	ACA Thr	GTT Val	GTA Val 300	GAT Asp	ACC Thr	TCT Ser	GAA Glu	GAA Glu 305	GAT Asp	AGT Ser	1149
45	TTT Phe	TCA Ser	TTA Leu 310	Cys	TTT Phe	TCT Ser	AAA Lys	TGT Cys 315	AGA Arg	ACA Thr	AAA Lys	AAT Asn	CTA Leu 320	CAA Gln	AAA Lys	GTA Val	1197
50			Ser					Lys					Ala			GAT Asp	1245
50		Сув					Asn					Lys				GTA Val 355	1293
55						Asn					Leu					GCA Ala	1341
60					Phe					Asp					Glu	GTT Val	1389

			TCT Ser 390														1437
5			GCC Ala														1485
10			AAT Asn														1533
15			GAT Asp														1581
20			AAA Lys														1629
25			GAA Glu 470														1677
			CAG Gln														1725
30			AAA Lys														1773
35			GCA Ala														1821
40			GAA Glu														1869
45			GAG Glu 550														1917
	Pro	Ala 565	ACC Thr	Thr	Thr	Gln	Asn 570	Ser	Val	Ala	Leu	Lys 575	Asn	Ala	Gly	Leu	1965
50		Ser	ACT Thr				Lys										2013
55	Asp	Glu	ACA Thr	Ser	Tyr 600	Lys	Gly	Lys	Lys	Ile 605	Pro	Lys	Asp	Gln	Lys 610	Ser	2061
60			ATT		Cys										Glu		2109
	CCA	CTT	ACA	TTT	GCA	AAT	GCT	GAT	TCA	GGT	TTA	TTG	CAT	TCT	TCT	GTG	2157

	Pro	Leu	Thr 630	Phe	Ala	Asn	Ala	Asp 635	Ser	Gly	Leu	Leu	His 640	Ser	Ser	Val	
5													ACT Thr				2205
10													AGA Arg				2253
15													TAT Tyr				2301
0.0	AAA Lys	TGT Cys	AAT Asn	AAG Lys 695	GAA Glu	AAA Lys	CTA Leu	CAG Gln	TTA Leu 700	TTT Phe	ATT Ile	ACC Thr	CCA Pro	GAA Glu 705	GCT Ala	GAT Asp	2349
20													GAT Asp 720				2397
25	AAA Lys	AAA Lys 725	GTT Val	TCA Ser	GAT Asp	ATA Ile	AAA Lys 730	GAA Glu	GAG Glu	GTC Val	TTG Leu	GCT Ala 735	GCA Ala	GCA Ala	TGT Cys	CAC His	2445
30													GAC Asp				2493
35													ACT Thr				2541
													ATG Met				2589
40									Asp				GGT Gly 800				2637
45			Asp										GAA Glu				2685
50		Val					Glu						GAG Glu			CCA Pro 835	2733
55						Arg					Ser		AAG Lys			Phe	2781
60					Asn					Gln					Glu	ACT Thr	2829
60																TTC Phe	2877

870 875 880

5	TCA Ser	GAC Asp 885	AAT Asn	GAG Glu	AAT Asn	AAT Asn	TTT Phe 890	GTC Val	TTC Phe	CAA Gln	GTA Val	GCT Ala 895	AAT Asn	GAA Glu	AGG Arg	AAT Asn	2925
10	AAT Asn 900	CTT Leu	GCT Ala	TTA Leu	GGA Gly	AAT Asn 905	ACT Thr	AAG Lys	GAA Glu	CTT Leu	CAT His 910	GAA Glu	ACA Thr	GAC Asp	TTG Leu	ACT Thr 915	2973
1.5				GAA Glu													3021
15				GAT Asp 935													3069
20				CTT Leu													3117
25				CTA Leu													3165
30				CCA Pro													3213
				CCA Pro					Ser					Phe			3261
35			Asn	AAG Lys 1015				Leu					Ile				3309
40		Met		TTC Phe			Ile					Pro					3357
45	Cys			ATT Ile		Asn					Asp					CTG Leu	3405
50		Lys			Ser		Asn			Ser					Ser	AGT Ser 1075	3453
				Ser					Ser		Ile					TTA Leu	3501
55					Asp					His					Ser	CAA Gln	3549
60				Ile			Leu		Thr					Ser		AGT Ser	3597

5	CAG TTT GAA TTT ACT CAG TTT AGA AAG CCA AGC TAC ATA TTG CAG AAG Gln Phe Glu Phe Thr Gln Phe Arg Lys Pro Ser Tyr Ile Leu Gln Lys 1125 1130 1135	3645
10	AGT ACA TTT GAA GTG CCT GAA AAC CAG ATG ACT ATC TTA AAG ACC ACT Ser Thr Phe Glu Val Pro Glu Asn Gln Met Thr Ile Leu Lys Thr Thr 1140 1145 1150 1155	3693
10	TCT GAG GAA TGC AGA GAT GCT GAT CTT CAT GTC ATA ATG AAT GCC CCA Ser Glu Glu Cys Arg Asp Ala Asp Leu His Val Ile Met Asn Ala Pro 1160 1165 1170	3741
15	TCG ATT GGT CAG GTA GAC AGC AGC AAG CAA TTT GAA GGT ACA GTT GAA Ser Ile Gly Gln Val Asp Ser Ser Lys Gln Phe Glu Gly Thr Val Glu 1175 1180 1185	3789
20	ATT AAA CGG AAG TTT GCT GGC CTG TTG AAA AAT GAC TGT AAC AAA AGT Ile Lys Arg Lys Phe Ala Gly Leu Leu Lys Asn Asp Cys Asn Lys Ser 1190 1195 1200	3837
25	GCT TCT GGT TAT TTA ACA GAT GAA AAT GAA GTG GGG TTT AGG GGC TTT Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe Arg Gly Phe 1205 1210 1215	3885
2.0	TAT TCT GCT CAT GGC ACA AAA CTG AAT GTT TCT ACT GAA GCT CTG CAA Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu Ala Leu Gln 1220 1225 1230 1235	3933
30	AAA GCT GTG AAA CTG TTT AGT GAT ATT GAG AAT ATT AGT GAG GAA ACT Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser Glu Glu Thr 1240 1245 1250	3981
35	TCT GCA GAG GTA CAT CCA ATA AGT TTA TCT TCA AGT AAA TGT CAT GAT Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Lys Cys His Asp 1255 1260 1265	4029
40	TCT GTT GTT TCA ATG TTT AAG ATA GAA AAT CAT AAT GAT AAA ACT GTA Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp Lys Thr Val 1270 1275 1280	4077
45	AGT GAA AAA AAT AAA TGC CAA CTG ATA TTA CAA AAT AAT ATT GAA Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn Asn Ile Glu 1285 1290 1295	4125
50	ATG ACT ACT GGC ACT TTT GTT GAA GAA ATT ACT GAA AAT TAC AAG AGA Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn Tyr Lys Arg 1300 1305 1310 1315	4173
50	AAT ACT GAA AAT GAA GAT AAC AAA TAT ACT GCT GCC AGT AGA AAT TCT Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser Arg Asn Ser 1320 1325 1330	4221
55	CAT AAC TTA GAA TTT GAT GGC AGT GAT TCA AGT AAA AAT GAT ACT GTT His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn Asp Thr Val 1335 1340 1345	4269
60	TGT ATT CAT AAA GAT GAA ACG GAC TTG CTA TTT ACT GAT CAG CAC AAC Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp Gln His Asn 1350 1355 1360	4317

5		Leu Lys	Leu Ser				GGA AAC ACT Gly Asn Thr	
J					Thr Phe		GTT GCG AAA Val Ala Lys	
10		Ala Cys					CAG TTA ACT Gln Leu Thr 1410	
15			Gln Asn	Ile Lys			TCT GAT ACA Ser Asp Thr 1425	
20						Val Ala	AAA GAG TCA Lys Glu Ser 440	
25		Ile Val	Asn Phe				GAA TTG CAT Glu Leu His	
					Ser Asp		AAG AAC AAA Lys Asn Lys	
30		Leu Ser					CAC AAA ATA His Lys Ile 1490	
35			Pro Val	Gly Thr			GTG ACC TTC Val Thr Phe 1505	
40						Glu Pro	ACT CTG TTG Thr Leu Leu 520	
45	Phe His 1525	Thr Ala	Ser Gly	Lys Lys 1530	Val Lys	Ile Ala 1535	AAG GAA TCT Lys Glu Ser	Leu
	Asp Lys 1540	: Val Lys	Asn Leu 1545	Phe Asp	Glu Lys	Glu Gln 1550		Glu 1555
50							AAG TAC AGA Lys Tyr Arg 1570	Glu
55			Leu Glu	Leu Ala			GAG ATC ACA Glu Ile Thr 1585	
60						Leu Asn	AAT GAT AAA Asn Asp Lys .600	
	CTT GTT	TCT ATI	GAG ACT	GTG GTG	CCA CCT	AAG CTC	TTA AGT GAT	AAT 5085

	Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu Ser Asp Asn 1605 1610 1615	
5	TTA TGT AGA CAA ACT GAA AAT CTC AAA ACA TCA AAA AGT ATC TTT TTG Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser Ile Phe Leu 1620 1625 1630 1635	5133
10	AAA GTT AAA GTA CAT GAA AAT GTA GAA AAA GAA ACA GCA AAA AGT CCT Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala Lys Ser Pro 1640 1645 1650	5181
15	GCA ACT TGT TAC ACA AAT CAG TCC CCT TAT TCA GTC ATT GAA AAT TCA Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile Glu Asn Ser 1655 1660 1665	5229
20	GCC TTA GCT TTT TAC ACA AGT TGT AGT AGA AAA ACT TCT GTG AGT CAG Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser Val Ser Gln 1670 1675 1680	5277
20	ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT GAT Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe Asp 1685 1690 1695	5325
25	GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT TTG Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr Leu 1700 1705 1710 1715	5373
30	TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT CAT Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn His 1720 1725 1730	5421
35	CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT AAC Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser Met Ser Asn 1735 1740 1745	5469
40	AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT CTC Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr Leu 1750 1760	5517
40	TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT GTT Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn Val 1765 1770 1775	5565
45	GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys 1780 1785 1790 1795	5613
50	GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu 1800 1805 1810	5661
55	GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT Glu Leu Val Thr Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile 1815 1820 1825	5709
60	AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala 1830 1835 1840	5757
	TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glu Thr Ile	5805

1845 1850 1855

5	AAA A Lys I 1860				Asp					Ser					Ile	5853
10	GAA A Glu A			Glu					Ile					Ile		5901
15	GGT T		Tyr					Asp					Leu			5949
10	CTA ( Leu A	qaA					Ser					Lys				5997
20	ATT ( Ile (					Ile					Gln					6045
25	GAG A Glu I 1940				Lys					Asp					Thr	6093
30	GAT A			Lys					Lys					Val		6141
2.5	GCA A		Thr					Ser					Lys			6189
35	GTA 1	Ser					Gln					Val				6237
40	GAA ( Glu A					Gln					Val					6285
45	GAA G Glu H 2020				Gln					Glu					Arg	6333
50	CCA (			Leu					Gly					Val		6381
r r	TCA S		Ala					Ser					Lys			6429
55	ATT :	Leu					His					Val				6477
60	GAT T Asp 1					Glu					Tyr					6525

5	CAA AAT GTA TCA AAA ATA CTT CCT CGT GTT GAT AAG AGA AAC CCA GAG Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg Asn Pro Glu 2100 2105 2110 2115	6573
10	CAC TGT GTA AAC TCA GAA ATG GAA AAA ACC TGC AGT AAA GAA TTT AAA His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys Glu Phe Lys 2120 2125 2130	6621
10	TTA TCA AAT AAC TTA AAT GTT GAA GGT GGT TCT TCA GAA AAT AAT CAC Leu Ser Asn Asn Leu Asn Val Glu Gly Gly Ser Ser Glu Asn Asn His 2135 2140 2145	6669
15	TCT ATT AAA GTT TCT CCA TAT CTC TCT CAA TTT CAA CAA GAC AAA CAA Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln Asp Lys Gln 2150 2155 2160	6717
20	CAG TTG GTA TTA GGA ACC AAA GTC TCA CTT GTT GAG AAC ATT CAT GTT Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn Ile His Val 2165 2170 2175	6765
25	TTG GGA AAA GAA CAG GCT TCA CCT AAA AAC GTA AAA ATG GAA ATT GGT Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met Glu Ile Gly 2180 2185 2190 2195	6813
2.0	AAA ACT GAA ACT TTT TCT GAT GTT CCT GTG AAA ACA AAT ATA GAA GTT Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn Ile Glu Val 2200 2205 2210	6861
30	TGT TCT ACT TAC TCC AAA GAT TCA GAA AAC TAC TTT GAA ACA GAA GCA Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu Thr Glu Ala 2215 2220 2225	6909
35	GTA GAA ATT GCT AAA GCT TTT ATG GAA GAT GAT GAA CTG ACA GAT TCT Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu Thr Asp Ser 2230 2235 2240	6957
40	AAA CTG CCA AGT CAT GCC ACA CAT TCT CTT TTT ACA TGT CCC GAA AAT Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys Pro Glu Asn 2245 2250 2255	7005
45	GAG GAA ATG GTT TTG TCA AAT TCA AGA ATT GGA AAA AGA AGA GGA GAG Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg Arg Gly Glu 2260 2265 2270 2275	7053
E O	CCC CTT ATC TTA GTG GGA GAA CCC TCA ATC AAA AGA AAC TTA TTA AAT Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn Leu Leu Asn 2280 2285 2290	7101
50	GAA TTT GAC AGG ATA ATA GAA AAT CAA GAA AAA TCC TTA AAG GCT TCA Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu Lys Ala Ser 2295 2300 2305	7149
55	AAA AGC ACT CCA GAT GGC ACA ATA AAA GAT CGA AGA TTG TTT ATG CAT Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu Phe Met His 2310 2315 2320	7197
60	CAT GTT TCT TTA GAG CCG ATT ACC TGT GTA CCC TTT CGC ACA ACT AAG His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg Thr Thr Lys 2325 2330 2335	7245

5					Ile					Phe					CAA Gln 2		7293
				Lys					Glu					Glu	AAA Lys 2370		7341
10			Asn					Gly					Gln		TCT Ser		7389
15		Arg					Arg					Thr			CCA Pro		7437
20	Lys					Pro					Ser				AGA Arg		7485
25					Arg					Glu					AAG Lys		7533
				Gly					Asp					Ile	AAT Asn 2450		7581
30			Ile					Lys					Gln		GCA Ala		7629
35		Thr					Glu					Asp			ACA Thr		7677
40	Leu					Asp					Arg				AAA Lys		7725
45					Phe					Ser					AAA Lys		7773
				Pro					Lys					Gly	CAA Gln 2530		7821
50			Ala					Gln					Gly		TCT Ser		7869
55		Cys					Ser					Ser			TTT Phe		7917
60	Thr					Gly					Trp				GGA Gly		7965
	CAG	TTG	GCT	GAT	GGT	GGA	TGG	CTC	ATA	CCC	TCC	AAT	GAT	GGA	AAG	GCT	8013

	Gln 2580	Leu	Ala	Asp	_	Gly 2585	Trp	Leu	Ile		Ser 2590	Asn	Asp	Gly	Lys 2	Ala 2595	
5				Glu					Leu					Gly	GTG Val 2610		8061
10			Leu					${\tt Trp}$					Tyr		TGG Trp		8109
15		Trp					Met					Pro			TTT Phe		8157
20	Asn					Pro					Leu				TAC Tyr		8205
20					Ile					Arg					AAG Lys		8253
25				Asp					Lys					Cys	GTT Val 2690		8301
30			Ile					Asn					Ser		AAT Asn		8349
35		Ser					Gln					Ile			ACA Thr		8397
40	Gly					Lys					Pro				GCT Ala		8445
10					Arg					Gln					CAT His		8493
45				Val					Ala					Glu	GCC Ala 2770		8541
50			Leu					Ser					Arg		GCT Ala		8589
55		Tyr					Phe					Arg			CCT Pro		8637
60	Pro					Phe					Asn				GTT Val		8685
<b>3 3</b>															ACA Thr		8733

3070

Phe Leu Asp Pro Asp Phe Gln Pro Ser Cys Ser Glu Val Asp Leu Ile

3065

3060

5				GTT Val					Lys					Pro		_	9501
10			Ser	GAC Asp 3095				Asn					Lys				9549
		Leu		GAG Glu			Ile					Leu					9597
15	Asn			TGG Trp		Pro					Gly						9645
20				TTT Phe	Ser					Ser					His		9693
25				TTC Phe					Asn					Ile			9741
30			Asn	GAA Glu 3175				Lys					Leu				9789
30		Pro		TGG Trp			Pro					Thr					9837
35	Thr			ATC Ile		Pro					Lys						9885
40				GAG Glu	Ile					Pro					Met		9933
45				TCT Ser					Val					Thr			9981
50			Lys	GGG Gly 3255				Ile					Asn				10029
		Arg		TTG Leu			Leu					Leu					10077
55	Ser			TGT Cys		Phe					Ala						10125
60		Pro		AGT Ser	Cys					Glu					Lys		10173

5	GAA CTG AAT TCT CCT CAG ATG ACT CCA TTT AAA AAA TTC AAT GAA ATT 1 Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe Asn Glu Ile 3320 3325 3330	.0221
-	TCT CTT TTG GAA AGT AAT TCA ATA GCT GAC GAA GAA CTT GCA TTG ATA 1 Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu Ala Leu Ile 3335 3340 3345	.0269
10	AAT ACC CAA GCT CTT TTG TCT GGT TCA ACA GGA GAA AAA CAA TTT ATA 1 Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys Gln Phe Ile 3350 3360	.0317
15	TCT GTC AGT GAA TCC ACT AGG ACT GCT CCC ACC AGT TCA GAA GAT TAT 1 Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser Glu Asp Tyr 3365 3370 3375	.0365
20	CTC AGA CTG AAA CGA CGT TGT ACT ACA TCT CTG ATC AAA GAA CAG GAG 1 Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys Glu Gln Glu 3380 3385 3390 3395	.0413
25	AGT TCC CAG GCC AGT ACG GAA GAA TGT GAG AAA AAT AAG CAG GAC ACA 1 Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys Gln Asp Thr 3400 3405 3410	0461
23	ATT ACA ACT AAA AAA TAT ATC TAA Ile Thr Thr Lys Lys Tyr Ile 3415	10485
30	(2) INFORMATION FOR SEQ ID NO:11:	
	(i) CROVENOR GUARAGERICA	
35	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 3418 amino acids</li><li>(B) TYPE: amino acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
35	<ul> <li>(A) LENGTH: 3418 amino acids</li> <li>(B) TYPE: amino acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> <li>(ii) MOLECULE TYPE: protein</li> <li>(v) FRAGMENT TYPE: internal</li> </ul>	
	<ul> <li>(A) LENGTH: 3418 amino acids</li> <li>(B) TYPE: amino acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> <li>(ii) MOLECULE TYPE: protein</li> <li>(v) FRAGMENT TYPE: internal</li> <li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:</li> </ul>	
	(A) LENGTH: 3418 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear  (ii) MOLECULE TYPE: protein (v) FRAGMENT TYPE: internal  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:  Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys 1 5 10 15	
40	(A) LENGTH: 3418 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear  (ii) MOLECULE TYPE: protein (v) FRAGMENT TYPE: internal  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:  Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys 1 5 10 15  Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe 20 25 30	
<b>4</b> 0 <b>4</b> 5	(A) LENGTH: 3418 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear  (ii) MOLECULE TYPE: protein (v) FRAGMENT TYPE: internal  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:  Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys 1	
40	(A) LENGTH: 3418 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear  (ii) MOLECULE TYPE: protein (v) FRAGMENT TYPE: internal  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:  Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys 1 5 10 15  Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe 20 25 30  Glu Glu Leu Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu 35 40 45  Glu Ser Glu His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr 50 55 60	
<b>4</b> 0 <b>4</b> 5	(A) LENGTH: 3418 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear  (ii) MOLECULE TYPE: protein (v) FRAGMENT TYPE: internal  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:  Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys 1 5 10  Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe 20 25 30  Glu Glu Leu Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu 35 40 45  Glu Ser Glu His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr 50 55 60  Pro Gln Arg Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile 65 70 75 80	
<b>4</b> 0 <b>4</b> 5	(A) LENGTH: 3418 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear  (ii) MOLECULE TYPE: protein (v) FRAGMENT TYPE: internal  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:  Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys 1 5 10 15  Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe 20 25 30  Glu Glu Leu Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu 35 40 45  Glu Ser Glu His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr 50 55 60  Pro Gln Arg Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile	
40 45 50	(A) LENGTH: 3418 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear  (ii) MOLECULE TYPE: protein (v) FRAGMENT TYPE: internal  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:  Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys 1 5 10 15  Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe 20 25 30  Glu Glu Leu Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu 35 40 45  Glu Ser Glu His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr 50 55 60  Pro Gln Arg Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile 65 70 75 80  Phe Lys Glu Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys 85 90 95  Glu Leu Asp Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser 100 105 110  Arg His Lys Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp	
40 45 50	(A) LENGTH: 3418 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear  (ii) MOLECULE TYPE: protein (v) FRAGMENT TYPE: internal  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:  Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys 1 5 10 15  Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe 20 25 30  Glu Glu Leu Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu 35 40 45  Glu Ser Glu His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr 50 55 60  Pro Gln Arg Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile 65 70 75 80  Phe Lys Glu Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys 85 90 95  Glu Leu Asp Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser	

	145					150					155					160
	-	-			165				_	170				Arg	175	
5	Pro	Lys	His	Ile 180	Ser	Glu	Ser	Leu	Gly 185	Ala	Glu	Val	Asp	Pro 190	Asp	Met
	Ser	Trp	Ser 195	Ser	Ser	Leu	Ala	Thr 200	Pro	Pro	Thr	Leu	Ser 205	Ser	Thr	Val
10	Leu	Ile 210	Val	Arg	Asn	Glu	Glu 215	Ala	Ser	Glu	Thr	Val 220	Phe	Pro	His	Asp
	Thr 225	Thr	Ala	Asn	Val	Lys 230	Ser	Tyr	Phe	Ser	Asn 235	His	Asp	Glu	Ser	Leu 240
	Lys	Lys	Asn	Asp	Arg 245	Phe	Ile	Ala	Ser	Val 250	Thr	Asp	Ser	Glu	Asn 255	Thr
15	Asn	Gln	Arg	Glu 260	Ala	Ala	Ser	His	Gly 265	Phe	Gly	Lys	Thr	Ser 270	Gly	Asn
			275				_	280					285	Ser		
20	Asn	Val 290	Leu	Glu	Asp	Glu	Val 295	Tyr	Glu	Thr	Val	Val 300	Asp	Thr	Ser	Glu
	305	_				310	_			_	315			Lys		320
		_			325		_			330				His	335	
25				340	_		_		345					Glu 350		
			355					360					365	Leu		
30		370					375					380		Lys		
	385					390					395			Leu		400
<u> </u>		_			405					410				Leu	415	
35			_	420					425	_	_			Asp 430		
		_	435	_	_	_		440					445	Leu		_
40		450				_	455		_			460		Glu		
	465		_	_	_	470					475			Thr		480
. –					485					490				Val	495	
45				500					505		_			Glu 510		
	-		515					520		_			525	Asp		
50		530	_				535				_	540		Ile		
	545					550					555			Ile		560
					565					570				Leu	575	
55				580				-	585					Phe 590		
			595					600					605	Pro		
60		610					615	_				620		Ala		
	625	GIU	ALd	PLO	ьeu	630	rne	Ald	ASII	AId	635	ser	GTÀ	Leu	ьеu	640

	Ser	Ser	Val	Lys	Arg 645	Ser	Cys	Ser	Gln	Asn 650	Asp	Ser	Glu	Glu	Pro 655	Thr
5	Leu	Ser	Leu	Thr 660		Ser	Phe	Gly	Thr 665		Leu	Arg	Lys	Cys 670		Arg
J	Asn	Glu	Thr 675		Ser	Asn	Asn	Thr 680		Ile	Ser	Gln	Asp 685		Asp	Tyr
	Lys	Glu 690		Lys	Cys	Asn	Lys 695		Lys	Leu	Gln	Leu 700		Ile	Thr	Pro
10	Glu 705	Ala	Asp	Ser	Leu	Ser 710	Cys	Leu	Gln	Glu	Gly 715	Gln	Cys	Glu	Asn	Asp 720
	Pro	Lys	Ser	Lys	Lys 725	Val	Ser	Asp	Ile	Lys 730	Glu	Glu	Val	Leu	Ala 735	Ala
15	Ala	Cys	His	Pro 740	Val	Gln	His	Ser	Lys 745	Val	Glu	Tyr	Ser	Asp 750	Thr	Asp
	Phe	Gln	Ser 755	Gln	Lys	Ser	Leu	Leu 760	Tyr	Asp	His	Glu	Asn 765	Ala	Ser	Thr
		Ile 770					775	_	_			780				
20	Ile 785	Ser	Arg	Gly	Lys	Glu 790	Ser	Tyr	Lys	Met	Ser 795	Asp	Lys	Leu	Lys	Gly 800
		Asn	_		805	_				810	_				815	
25		Asn		820		_			825			_	_	830		
		Leu	835			_	_	840	_				845		_	-
		Gln 850					855					860		_		
30	865	Glu				870		_			875			_		880
		Leu			885					890					895	
35		Arg		900				_	905		_			910		
		Leu	915					920			_		925			
40		Tyr 930	_	_		_	935	_				940				_
40	945	Asp				950					955					960
		His		_	965			_		970		_		_	975	
45		Asn Ala		980	_				985			_	-	990		-
		Arg	995					1000	)				1005	5	_	
50		1010 Lys	)				1015	5		-		1020	)			
50	102	5				1030	)				1035	5		_		104
		Leu			1045	5				1050	)				1055	5
55		Lys		1060	)				1065	5				1070	)	
		Ser	1075	5				1080	)				1085	5		
60		Met 1090	)				1099	5				1100	)			
	110					1110	)				1115	5				112
	ser	Gly	261	GIII	File	GIU	Pne	Inr	GIII	ьце	arg	ьys	Pro	ser	ıyr	тте

				1125					1130					1135	;
	Leu Gl	n Lys	Ser 1140		Phe	Glu		Pro 1145		Asn	Gln	Met	Thr 1150		Leu
5	Lys Th	1155	5			_	1160					1165	;		
	Asn Al	a Pro	Ser	Ile	Gly	Gln 1175		Asp	Ser	Ser	Lys 1180		Phe	Glu	Gly
10	Thr Va 1185			-	1190	)				1195					120
	Asn Ly			1205					1210	)				1215	5
	Arg Gl	•	1220				-	1225	5				1230	)	
15	Ala Le	1235	5				1240	+				1245	i		
		:50				1255	5				1260	)			
20	Cys Hi 1265	_			1270	)				1275	,				128
	Lys Th			1285	i				1290	)				1295	5
0.5	Asn Il		1300	)				1305	5				1310	)	
25	Tyr Ly	1315	5				1320	)				1325	5		
		30				1335	5				1340	)			
30	Asp Th		_		1350	)				1355	5				136
	Gln Hi			1365	5				1370	)				1375	5
	Asn Th		1380	)				1385	5				1390	)	
35	Ala Ly	139	5				1400	)				1405	5		
		nr Ala 110	Thr	Lys	Thr	Glu 1415		Asn	Ile	Lys	Asp 1420		Glu	Thr	Ser
40	Asp Th	nr Phe	Phe	Gln	Thr 1430		Ser	Gly	Lys	Asn 1435		Ser	Val	Ala	Lys 144
	Glu Se	er Phe	Asn	Lys 1445		Val	Asn	Phe	Phe 1450		Gln	Lys	Pro	Glu 145	
		is Asn	1460	)				1465	5				1470	)	
45	Asn Ly	ys Met 147!	_	Ile	Leu	Ser	Tyr 1480		Glu	Thr	Asp	Ile 1489		Lys	His
	14	le Leu 190				149	5				1500	0			
50	Thr Pl 1505	ne Gln	Gly	Gln	Pro 151		Arg	Asp	Glu	Lys 151		Lys	Glu	Pro	Thr 152
		eu Gly	Phe	His 1525	Thr		Ser	Gly	Lys 1530		Val	Lys	Ile	Ala 153	
	Glu Se	er Leu	Asp 1540		Val	Lys	Asn	Leu 154		Asp	Glu	Lys	Glu 155		Gly
55		er Glu 155	5				1560	)		-		156	5		
	1!	rg Glu 570				157	5				158	0			
60	1585	hr Ala			159	0				159	5				160
	Asp L	ys Asn	Leu	Val 160		Ile	Glu	Thr	Val 161		Pro	Pro	Lys	Leu 161	

	Ser	Asp	Asn	Leu 1620		Arg	Gln	Thr	Glu 1625		Leu	Lys	Thr	Ser 1630		Ser
5	Ile	Phe	Leu 1635		Val	Lys		His 1640		Asn	Val		Lys 1645		Thr	Ala
	Lys	Ser 1650		Ala	Thr	Сув	Tyr 1655		Asn	Gln	Ser	Pro 1660		Ser	Val	Ile
	Glu 1669	Asn	Ser	Ala	Leu	Ala 1670		Tyr	Thr	Ser	Cys 1675		Arg	Lys	Thr	Ser 168
10		Ser	Gln	Thr	Ser 1685		Leu	Glu	Ala	Lys 1690		Trp	Leu	Arg	Glu 1695	
	Ile	Phe	Asp	Gly 1700		Pro	Glu	Arg	Ile 1705		Thr	Ala	Asp	Tyr 1710		Gly
15		Tyr	1715	5				1720	)				1725	5		
	Lys	Asn 1730		Leu	Ser	Glu	Lys 1735		Asp	Thr	Tyr	Leu 1740		Asn	Ser	Ser
	174					1750	)				1755	5				176
20	Gly	Tyr	Leu	Ser	Lys 1765		Lys	Leu	Asp	Ser 1770		Ile	Glu	Pro	Val 1775	
	Lys	Asn	Val	Glu 1780	_	Gln	Lys	Asn	Thr 1789		Phe	Ser	Lys	Val 1790		Ser
25		Val	1795	5				1800	)				1809	5		
	-	Val 1810	)				1819	5				1820	)			
	182			-		1830	)				1835	5				184
30		Pro			1845	5			_	1850	)				1855	5
		Thr		1860	)		_		186	5				1870	)	
35		Ile	1879	5				1880	ַ כ כ		_		188	5		
		Met 1890	)	_	_		189	5				1900	)			
4.0	190					1910	0				1915	5				192
40		Ala	_		1925	5				1930	)				1935	5
		Gly		1940	ר - ס			-	194	5				1950	)	
45		Thr	195	5		_	-	1960	0				196	5		
		Ser 1970	)				197	5				198	0			
50	198	Val 5 Glu				199	0				199	5				200
30		Ser			2009	5				2010	C				2015	5
	_	Arg		2020	0				202	5				2030	0	
55		Val	203	5				204	0		_		204	5		
		205 Val	0				205	5				206	0			
60	206					207	0				207	5				208
<b>-</b>		Ser		_	208	5				209	0				209	5
			_					-				_		_		

				2100	1				2105	;				2110		
			2115	-				2120	)				2125	,		
5		2130	)	Leu			2135	i				2140	)			
	2145	5		Ser		2150	)				2155	,				216
10	-	-		Gln	2165	5				2170	)				2175	5
				Leu 2180	)	-			2185	5				2190	)	
<b>.</b>			2195					2200	)				2205	5		
15		2210	)	Cys			2215	5				2220	)			
	2225	5		Val Lys		2230	)	_			2235	5				224
20				Glu	2245	5				2250	)				2255	5
				2260 Pro	)				2265	5				2270	)	
25	_	-	2275					2280	)				2285	5		
		2290	)	Lys			2295	5				2300	)			
	230	5		His		2310	)				2315	5				232
30				Glu	2325	5				2330	)				2335	5
	Gly	Gln		2340 Phe		Ser	Lys				Tyr	Glu				Leu
35	Glu	-		Ser	Ser	Asn				Ser	Gly	His 2380			Tyr	Gln
	Val 238			Thr	Arg	Asn 2390			Met	Arg	His 2395	Leu		Thr	Thr	Gly 240
40			Thr	Lys	Val 2409	Phe		Pro	Pro	Phe 2410	Lys		Lys	Ser	His 2419	Phe
- 4	His	Arg	Val	Glu 2420	Gln		Val	Arg	Asn 2425	Ile		Leu	Glu	Glu 2430	Asn	
	Gln	Lys	Gln 243	Asn		Asp	Gly	His 2440	_	Ser	Asp	Asp	Ser 244		Asn	Lys
45	Ile	Asn 245	_	Asn	Glu	Ile	His 245		Phe	Asn	Lys	Asn 2460		Ser	Asn	Gln
	246	5		Val		2470	С				2475	5				248
50				Leu	2485	5				2490	)				249	5
	_	_		Arg 2500	)				250	5				2510	)	
c c		_	251					2520	)				252	5		
55	_	253	0	Pro His			253	5				254	C			
	254	5		Thr		255	0				255	5				256
60				Gln	256	5	_			2570	0				257	5
	-1-	1		2580				1	258					259		- 1-

	Gly	Lys	Ala 2599		Lys	Glu	Glu	Phe 2600		Arg	Ala	Leu	Cys 2609		Thr	Pro
	Gly	Val	Asp	Pro	Lys	Leu	Ile	Ser	Arg	Ile	Trp	Val	Tyr	Asn	His	Tyr
5		2610					261					2620				
			Ile	Ile	Trp			Ala	Ala	Met			Ala	Phe	Pro	
	2625		ת 1 ת	Λαn	λrα	2630		Cor	Dro	Clu	263		T 011	Lou	CIn	264
	GIU	PIIC	Ата	ASII	Arg 2645		ьец	ser	PIO	2650	_	Val	ьеи	ьеи	2655	
10	Lys	Tyr	Arq	Tyr	Asp		Glu	Ile	Asp			Arq	Ara	Ser		
	•	1		2660					2669					2670		
	Lys	Lys	Ile	Met	Glu	Arg	Asp	Asp	Thr	Ala	Ala	Lys			Val	Let
		_	2675		_	_		2680					2685			
1.5	Cys			Asp	Ile	Ile			Ser	Ala	Asn			Glu	Thr	Ser
15	Car	2690		Thr	Ser	Sor	2695		Thr	@1 n	Tura	2700		т	Tla	<b>01.</b>
	2705		цуѕ	1111	SEI	2710		Asp	1111		2715		Ala	TTE	116	272
			Asp	Gly	Trp			Val	Lvs				Asp	Pro	Pro	
			-	•	2725				•	2730			-		2735	
20	Leu	Ala	Val		Lys	Asn	Gly	Arg	Leu	Thr	Val	Gly	Gln	Lys	Ile	Ile
	_	•		2740		_			2745		_			2750		
	Leu	His			Glu	Leu	Val			Pro	Asp	Ala			Pro	Leu
	Glu	Δla	2755		Ser	T.e11	Met	2760		Tla	Sar	λla	2765		Thr	7 20
25	Olu	2770		Olu	DCI	пси	2775		цуз	110	561	2780		261	1111	Arg
	Pro			Trp	Tyr	Thr			Gly	Phe	Phe			Pro	Arg	Pro
	2785	5				2790	)				2799	5				280
	Phe	Pro	Leu	Pro	Leu		Ser	Leu	Phe			Gly	Gly	Asn	Val	Gly
2.0	~	** 7	_		2805		~ 7	_		2810				_	2815	
30	Cys	vaı	Asp	Val 2820	Ile	IIe	GIn	_		_	Pro	Ile	Gin			Glu
	Lvs	Thr	Ser		Gly	T.e.ii	Туг		2825		Acn	Glu	Ara	2830		Gl 11
	2,5		2835		Gly	шси	- y -	2840		ALG	Veii	Giu	2845		GIU	Giu
	Lys	Glu	Ala	Ala	Lys	Tyr	Val			Gln	Gln	Lys			Glu	Ala
35		2850	)				2855	5				2860	)			
			Thr	Lys	Ile			Glu	Phe	Glu			Glu	Glu	Asn	
	2865		Dxo	TT	т от	2870		7 22.00	77.	T	2875		α1	<b>~</b> 11	**- 7	288
	TIII	пуъ	PIO	TÀT	Leu 2885		ser	Arg	Ala	2890		Arg	GIII	GIII	2895	_
40	Ala	Leu	Gln	Asp	Gly		Glu	Leu	Tyr			Val	Lvs	Asn		
				2900	)				2905	5				2910	)	
	Asp	Pro			Leu											
	<b>.</b>				_				-					•		
45	Leu	Asn 2930		His	Arg	GIn	Met 2935		Asn	Asp	Lys			Ala	Gln	Ile
10	Gln			Tle	Arg	Tivs			Glu	Ser	Δla	2940		Taye	G311	Gln
	2945		014		9	2950			Olu	001	2955		0111	цуз	Olu	296
	Gly	Leu	Ser	Arg	Asp	Val	Thr	Thr	Val	Trp			Arg	Ile	Val	
					2965					2970					2975	
50	Tyr	Ser	Lys		Glu	Lys	Asp	Ser			Leu	Ser	Ile		_	Pro
	Sar	Car	Λαn	2980		Cor	T 011	T 0.1	2985		<b>a</b> 1	T	7	2990		<b>-</b> 1 -
	ser	261	2995		Tyr	ser	ьеu	3000		GIU	GIY	ьуѕ	Arg		Arg	шe
	Tyr	His			Thr	Ser	Lvs			Ser	Lvs	Ser			Ala	Asn
55	-	3010					3015				_1	3020		5		
	Ile	Gln	Leu	Ala	Ala	Thr	Lys	Lys	Thr	Gln	Tyr	Gln	Gln	Leu	Pro	Val
	3025				_	3030					3035		_			304
	Ser	Asp	Glu	Пе	Leu		Gln	Ile	Tyr			Arg	Glu	Pro		
60	Phe	Ser	Lve	Phe	3045 Leu		Pro	Δen	Dhe	3050		Ser	Cve	Ser	3055	
	- 110	JC1	- y 5	3060		~25P	110	Asp	3065		FIU	PGT	Cyp	3070		val
	Asp	Leu	Tle		Phe	Val	Val	Ser			Laze	Laze	Thr			λla

	3075		3080	3085
		eu Ser Asp 3095	Glu Cys Tyr A	sn Leu Leu Ala Ile Lys 3100
5	Phe Trp Ile Asp Lo	eu Asn Glu 3110		ys Pro His Met Leu Ile 115 312
	Ala Ala Ser Asn L			er Lys Ser Gly Leu Leu 3135
10	Thr Leu Phe Ala G 3140	ly Asp Phe		er Ala Ser Pro Lys Glu 3150
	Gly His Phe Gln G 3155	lu Thr Phe	Asn Lys Met L 3160	ys Asn Thr Val Glu Asn 3165
	Ile Asp Ile Leu C	ys Asn Glu 3175		ys Leu Met His Ile Leu 3180
15	His Ala Asn Asp P: 3185	ro Lys Trp 3190		hr Lys Asp Cys Thr Ser 195 320
		la Gln Ile 205	Ile Pro Gly T 3210	hr Gly Asn Lys Leu Leu 3215
20	Met Ser Ser Pro A	sn Cys Glu	Ile Tyr Tyr G 3225	ln Ser Pro Leu Ser Leu 3230
	Cys Met Ala Lys A: 3235	rg Lys Ser	Val Ser Thr P 3240	ro Val Ser Ala Gln Met 3245
	Thr Ser Lys Ser C	ys Lys Gly 3255		le Asp Asp Gln Lys Asn 3260
25	3265	3270	3	er Arg Leu Pro Leu Pro 275 328
		ro Ile Cys 285	Thr Phe Val S 3290	er Pro Ala Ala Gln Lys 3295
30	Ala Phe Gln Pro P	ro Arg Ser	Cys Gly Thr L 3305	ys Tyr Glu Thr Pro Ile 3310
	3315		3320	hr Pro Phe Lys Lys Phe 3325
	3330	3335	i	le Ala Asp Glu Glu Leu 3340
35	3345	3350	3	ly Ser Thr Gly Glu Lys 355 336
	3:	365	3370	hr Ala Pro Thr Ser Ser 3375
40	3380		3385	hr Thr Ser Leu Ile Lys 3390
	3395		3400	lu Cys Glu Lys Asn Lys 3405
4.5	Gln Asp Thr Ile T	hr Thr Lys 3415		
45	(2) INFO	RMATION FOR	SEQ ID NO:12	:
	(i) SEQUENC			
50	(B) TYPE:	H: 10485 ba	id -	
		DEDNESS: do OGY: linear		
55	(ii) MOLECU: (ix) FEATUR:		NA	
J J		e: /KEY: Codin	ia Seguence	
	(B) LOCA	TION: 229		I5)
60			CION: SEQ ID N	·

5	GGTGGCGG TCTGCTGG ACAGATT CTGGAGCG	CGC CT IGT GA	CGGGTG1 CCGGCGC	C TT	TTGC TTTT	GGCG GTCA	GTC GCT	GGTC TACT	GCC CCCG	GCCA	GGAC AAAA	SAA C AAG A A ATC	CCTC ACTC CCT	SAGGGG SCACCT	60 120 180 237
10	GGA TCC Gly Ser 5			Pro '											285
15	AAC AAA Asn Lys 20														333
20	TCT TCA Ser Ser														381
2. 0	CAT AAA His Lys		sn Asn												429
25	AAA CCA Lys Pro														477
30	CAA GGG Gln Gly 85			Pro :											525
35	AAA TTC Lys Phe 100														573
40	AGT CTT Ser Leu														621
10	TGT CCA Cys Pro	Leu L													669
45	TGT ACA Cys Thr														717
50	TTG TTT Leu Phe 165			Lys :											765
55	ATT TCT Ile Ser 180														813
60	AGT TCT Ser Ser														861
50	AGA AAT Arg Asn														909

AAT GTG AAA AGC TAT TTT TCC AAT CAT GAT GAA AGT CTG AAG AAA AAT Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu Lys Lys Asn GAT AGA TTT ATC GCT TCT GTG ACA GAC AGT GAA AAC ACA AAT CAA AGA Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr Asn Gln Arg GAA GCT GCA AGT CAT GGA TTT GGA AAA ACA TCA GGG AAT TCA TTT AAA Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn Ser Phe Lys GTA AAT AGC TGC AAA GAC CAC ATT GGA AAG TCA ATG CCA CAT GTC CTA Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro His Val Leu GAA GAT GAA GTA TAT GAA ACA GTT GTA GAT ACC TCT GAA GAA GAT AGT Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu Glu Asp Ser TTT TCA TTA TGT TTT TCT AAA TGT AGA ACA AAA AAT CTA CAA AAA GTA Phe Ser Leu Cys Phe Ser Lys Cys Arg Thr Lys Asn Leu Gln Lys Val AGA ACT AGC AAG ACT AGG AAA AAA ATT TTC CAT GAA GCA AAC GCT GAT Arg Thr Ser Lys Thr Arg Lys Lys Ile Phe His Glu Ala Asn Ala Asp GAA TGT GAA AAA TCT AAA AAC CAA GTG AAA GAA AAA TAC TCA TTT GTA Glu Cys Glu Lys Ser Lys Asn Gln Val Lys Glu Lys Tyr Ser Phe Val TCT GAA GTG GAA CCA AAT GAT ACT GAT CCA TTA GAT TCA AAT GTA GCA Ser Glu Val Glu Pro Asn Asp Thr Asp Pro Leu Asp Ser Asn Val Ala CAT CAG AAG CCC TTT GAG AGT GGA AGT GAC AAA ATC TCC AAG GAA GTT His Gln Lys Pro Phe Glu Ser Gly Ser Asp Lys Ile Ser Lys Glu Val GTA CCG TCT TTG GCC TGT GAA TGG TCT CAA CTA ACC CTT TCA GGT CTA Val Pro Ser Leu Ala Cys Glu Trp Ser Gln Leu Thr Leu Ser Gly Leu AAT GGA GCC CAG ATG GAG AAA ATA CCC CTA TTG CAT ATT TCT TCA TGT Asn Gly Ala Gln Met Glu Lys Ile Pro Leu Leu His Ile Ser Ser Cys GAC CAA AAT ATT TCA GAA AAA GAC CTA TTA GAC ACA GAG AAC AAA AGA Asp Gln Asn Ile Ser Glu Lys Asp Leu Leu Asp Thr Glu Asn Lys Arg AAG AAA GAT TTT CTT ACT TCA GAG AAT TCT TTG CCA CGT ATT TCT AGC Lys Lys Asp Phe Leu Thr Ser Glu Asn Ser Leu Pro Arg Ile Ser Ser 

CTA CCA AAA TCG GAG AAG CCA TTA AAT GAG GAA ACA GTG GTA AAT AAG

Leu Pro Lys Ser Glu Lys Pro Leu Asn Glu Glu Thr Val Val Asn Lys

5		GAG Glu							1677
10		GCA Ala							1725
10		AAG Lys							1773
15		AGT Ser							1821
20		GCC Ala 535							1869
25		GAC Asp							1917
30		ACC Thr							1965
30		TTG Leu							2013
35		TCT Ser							2061
40		AAC Asn 615							2109
45		TTT Phe							2157
50		TGT Cys							2205
50		TTT Phe							2253
55		AAT Asn							2301
60		AAG Lys 695							2349

_			TCA Ser 710														2397
5			GTT Val														2445
10			CAA Gln														2493
15			AGT Ser														2541
20	Thr	Pro	ACT Thr	Ser 775	Lys	Asp	Val	Leu	Ser 780	Asn	Leu	Val	Met	Ile 785	Ser	Arg	2589
25			GAA Glu 790														2637
23	GAA Glu	TCT Ser 805	GAT Asp	GTT Val	GAA Glu	TTA Leu	ACC Thr 810	AAA Lys	AAT Asn	ATT Ile	CCC Pro	ATG Met 815	GAA Glu	AAG Lys	AAT Asn	CAA Gln	2685
30	GAT Asp 820	Val	TGT Cys	GCT Ala	TTA Leu	AAT Asn 825	GAA Glu	AAT Asn	TAT Tyr	AAA Lys	AAC Asn 830	GTT Val	GAG Glu	CTG Leu	TTG Leu	CCA Pro 835	2733
35			AAA Lys														2781
40	AAC Asn	CAA Gln	AAC Asn	ACA Thr 855	AAT Asn	CTA Leu	AGA Arg	GTA Val	ATC Ile 860	CAA Gln	AAA Lys	AAT Asn	CAA Gln	GAA Glu 865	GAA Glu	ACT Thr	2829
45	ACT Thr	TCA Ser	ATT Ile 870	TCA Ser	AAA Lys	ATA Ile	ACT Thr	GTC Val 875	AAT Asn	CCA Pro	GAC Asp	TCT Ser	GAA Glu 880	GAA Glu	CTT Leu	TTC Phe	2877
			AAT Asn					Val					Asn				2925
50		Leu	GCT Ala				Thr					Glu					2973
55						Ile					Thr					GGA Gly	3021
60					Lys					Val					Asp	TTG Leu	3069
	GTT	TAT	GTT	CTT	GCA	GAG	GAG	AAC	: AAA	raa .	AGT	GTA	AAG	CAG	CAT	ATA	3117

	Val	Tyr	Val 950	Leu	Ala	Glu	Glu	Asn 955	Lys	Asn	Ser	Val	Lys 960	Gln	His	Ile	
5		ATG Met 965															3165
10		AAA Lys															3213
15		TTA Leu		Pro					Ser					Phe			3261
20		TCA Ser	Asn					Leu					Ile				3309
20		ATG Met					Ile					Pro				_	3357
25	Cys	GTT Val 1045				Asn					Asp						3405
30		AAG Lys			Ser					Ser					Ser		3453
35		GTT Val		Ser					Ser					Gln			3501
40		TCC Ser	Lys					Ser					Thr				3549
40		GCA Ala					Leu					Glu					3597
45	Gln	TTT Phe				Gln					Ser						3645
50		ACA Thr			Val					Met					Thr		3693
55		GAG Glu		Cys					Leu					Asn			3741
60		ATT	Gly														3789
30		AAA Lys															3837

1190 1195 1200

5	GCT TCT GGT TAT TTA ACA Ala Ser Gly Tyr Leu Thr 1205	GAT GAA AAT GAA Asp Glu Asn Glu 210	GTG GGG TTT AGG GGC TTT Val Gly Phe Arg Gly Phe 1215	3885
10	TAT TCT GCT CAT GGC ACA Tyr Ser Ala His Gly Thr 1220 1225	Lys Leu Asn Val		3933
15	AAA GCT GTG AAA CTG TTT Lys Ala Val Lys Leu Phe 1240			3981
	TCT GCA GAG GTA CAT CCA Ser Ala Glu Val His Pro 1255			4029
20	TCT GTT GTT TCA ATG TTT Ser Val Val Ser Met Phe 1270			4077
25	AGT GAA AAA AAT AAT AAA Ser Glu Lys Asn Asn Lys 1285			4125
30	ATG ACT ACT GGC ACT TTT Met Thr Thr Gly Thr Phe 1300 1305	Val Glu Glu Ile 1	Thr Glu Asn Tyr Lys Arg	4173
35	AAT ACT GAA AAT GAA GAT Asn Thr Glu Asn Glu Asp 1320			4221
	CAT AAC TTA GAA TTT GAT His Asn Leu Glu Phe Asp 1335		AGT AAA AAT GAT ACT GTT Ser Lys Asn Asp Thr Val 1345	4269
40			TTT ACT GAT CAG CAC AAC Phe Thr Asp Gln His Asn 1360	4317
45	Ile Cys Leu Lys Leu Ser		AAG GAG GGA AAC ACT CAG Lys Glu Gly Asn Thr Gln 1375	4365
50		Asp Leu Thr Phe	TTG GAA GTT GCG AAA GCT Leu Glu Val Ala Lys Ala 1390 1395	4413
55			AAA GAA CAG TTA ACT GCT Lys Glu Gln Leu Thr Ala 1410	4461
			GAG ACT TCT GAT ACA TTT Glu Thr Ser Asp Thr Phe 1425	4509
60			GTC GCC AAA GAG TCA TTT Val Ala Lys Glu Ser Phe 1440	4557

5	Asn					Phe					Pro				CAT His	4605
10					Ser					Asp					AAA Lys	4653
10				Ser					Asp					Lys	ATA Ile 1490	4701
15			Ser					Thr					Val		TTC Phe	4749
20		Gln					Glu					Pro			TTG Leu	4797
25	Phe					Gly					Ile				TCT Ser	4845
30					Asn					Lys					AGT Ser	4893
30				Phe					Ala					Tyr	AGA Arg 1570	4941
35			Lys					Ala					Glu		ACA Thr	4989
40		Pro					Met					Asn			AAA Lys	5037
45	Leu					Thr					Lys				GAT Asp	5085
50		Cys			Thr					Thr					TTT Phe	5133
30				Val					Glu					Lys	AGT Ser 1650	5181
55			Cys					Ser					Ile		AAT Asn	5229
60		Leu					Ser					Thr			AGT Ser	5277

5	ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT GAT Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe Asp 1685 1690 1695	5325
5	GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT TTG Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr Leu 1700 1715	5373
10	TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT CAT Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn His 1720 1725 1730	5421
15	CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT AAC Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser Met Ser Asn 1735 1740 1745	5469
20	AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT CTC  Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr Leu  1750 1755 1760	5517
25	TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT GTT Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn Val 1765 1770 1775	5565
23	GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys 1780 1785 1790 1795	5613
30	GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu 1800 1805 1810	5661
35	GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT Glu Leu Val Thr Ser Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile 1815 1820 1825	5709
40	AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala 1830 1835 1840	5757
45	TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glu Thr Ile 1845 1850 1855	5805
10	AAA AAA GTG AAA GAC ATA TTT ACA GAC AGT TTC AGT AAA GTA ATT AAG Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys 1860 1865 1870 1875	5853
50	GAA AAC AAC GAG AAT AAA TCA AAA ATT TGC CAA ACG AAA ATT ATG GCA Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys Ile Met Ala 1880 1885 1890	5901
55	GGT TGT TAC GAG GCA TTG GAT GAT TCA GAG GAT ATT CTT CAT AAC TCT Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu His Asn Ser 1895 1900 1905	5949
60	CTA GAT AAT GAT GAA TGT AGC ACG CAT TCA CAT AAG GTT TTT GCT GAC Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val Phe Ala Asp 1910 1915 1920	5997
	ATT CAG AGT GAA GAA ATT TTA CAA CAT AAC CAA AAT ATG TCT GGA TTG	6045

	Ile Gln 1925	Ser Glu	Glu Ile	Leu 1930	Gln His	s Asn	Gln As		Ser Gly	Leu	
5	GAG AAA Glu Lys 1940	GTT TCT Val Ser	AAA ATA Lys Ile 1945	e Ser	CCT TG	s Asp	GTT AG Val So 1950	GT TTG er Leu	Glu Thr	TCA Ser .955	6093
10	GAT ATA Asp Ile	Cys Lys									6141
15	GCA AAT Ala Asn	ACT TGT Thr Cys 1975	GGG AT	TTTT Phe	AGC AC Ser Th	r Ala	AGT G Ser G	ly Lys	TCT GTC Ser Val 985	CAG Gln	6189
20	GTA TCA Val Ser	GAT GCT Asp Ala 1990	TCA TT	ı Gln	AAC GC Asn Al 1995	A AGA a Arg	CAA G Gln V	TG TTT al Phe 2000	TCT GAA Ser Glu	ATA Ile	6237
20	GAA GAT Glu Asp 2005						Val L				6285
25	GAA CAT Glu His 2020	TCA GAC Ser Asp	CAG CT Gln Le	u Thr	AGA GA Arg Gl	u Glu	AAT A Asn T 2030	ACT GCT Thr Ala	Ile Arg	ACT Thr 2035	6333
30	CCA GAA Pro Glu	CAT TTA His Leu	ATA TC Ile Se 2040	C CAA r Gln	AAA GG Lys Gl	C TTT y Phe 2045	TCA T Ser T	CAT AAT Cyr Asn	GTG GTA Val Val 2050	AAT Asn	6381
35	TCA TCT Ser Ser	GCT TTC Ala Phe 2055	Ser Gl	A TTT y Phe	AGT AC Ser Th 206	r Ala	AGT G	Sly Lys	CAA GTT Gln Val 2065	TCC Ser	6429
40	Ile Leu	GAA AGT Glu Ser 2070	TCC TT Ser Le	u His	AAA GT Lys Va 2075	T AAG 1 Lys	GGA G	GTG TTA /al Leu 2080	GAG GAA Glu Glu	TTT Phe	6477
40	GAT TTA Asp Leu 2085	ATC AGA	ACT GA Thr Gl	G CAT u His 2090	Ser Le	T CAC	Tyr S	CCA CCT Ser Pro	ACG TCT Thr Ser	AGA Arg	6525
45				e Leu		g Val			AAC CCA Asn Pro		6573
50	CAC TGT His Cys	GTA AAC Val Asr	C TCA GA 1 Ser Gl 2120	A ATG u Met	GAA AA Glu Ly	A ACC s Thr 2125	Cys S	AGT AAA Ser Lys	GAA TTT Glu Phe 2130	AAA Lys	6621
55			n Leu As			y Gly		Ser Glu	AAT AAT Asn Asn 2145		6669
60	Ser Ile	AAA GT Lys Val 2150	TCT CC L Ser Pr	A TAI	CTC TO Leu Se 2155	CT CAA	TTT (	CAA CAA Gln Gln 2160	GAC AAA Asp Lys	CAA Gln	6717
00									ATT CAT		6765

2165 2170 2175

5	TTG GGA AAA GAA CAG GCT TCA CCT AAA AAC GTA AAA ATG GAA ATT GGT Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met Glu Ile Gly 2180 2185 2190 2195	6813
10	AAA ACT GAA ACT TTT TCT GAT GTT CCT GTG AAA ACA AAT ATA GAA GTT Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn Ile Glu Val 2200 2205 2210	6861
1.5	TGT TCT ACT TAC TCC AAA GAT TCA GAA AAC TAC TTT GAA ACA GAA GCA Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu Thr Glu Ala 2215 2220 2225	6909
15	GTA GAA ATT GCT AAA GCT TTT ATG GAA GAT GAT GAA CTG ACA GAT TCT Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu Thr Asp Ser 2230 2235 2240	6957
20	AAA CTG CCA AGT CAT GCC ACA CAT TCT CTT TTT ACA TGT CCC GAA AAT Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys Pro Glu Asn 2245 2250 2255	7005
25	GAG GAA ATG GTT TTG TCA AAT TCA AGA ATT GGA AAA AGA AGA GGA GAG Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg Arg Gly Glu 2260 2265 2270 2275	7053
30	CCC CTT ATC TTA GTG GGA GAA CCC TCA ATC AAA AGA AAC TTA TTA AAT Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn Leu Leu Asn 2280 2285 2290	7101
	GAA TTT GAC AGG ATA ATA GAA AAT CAA GAA AAA TCC TTA AAG GCT TCA Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu Lys Ala Ser 2295 2300 2305	7149
35	AAA AGC ACT CCA GAT GGC ACA ATA AAA GAT CGA AGA TTG TTT ATG CAT Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu Phe Met His 2310 2315 2320	7197
40	CAT GTT TCT TTA GAG CCG ATT ACC TGT GTA CCC TTT CGC ACA ACT AAG His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg Thr Thr Lys 2325 2330 2335	7245
45	GAA CGT CAA GAG ATA CAG AAT CCA AAT TTT ACC GCA CCT GGT CAA GAA Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro Gly Gln Glu 2340 2345 2350 2355	7293
50	TTT CTG TCT AAA TCT CAT TTG TAT GAA CAT CTG ACT TTG GAA AAA TCT Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu Glu Lys Ser 2360 2365 2370	7341
55	TCA AGC AAT TTA GCA GTT TCA GGA CAT CCA TTT TAT CAA GTT TCT GCT Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln Val Ser Ala 2375 2380 2385	7389
ب ر	ACA AGA AAT GAA AAA ATG AGA CAC TTG ATT ACT ACA GGC AGA CCA ACC Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly Arg Pro Thr 2390 2395 2400	7437
60	AAA GTC TTT GTT CCA CCT TTT AAA ACT AAA TCA CAT TTT CAC AGA GTT Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe His Arg Val 2405 2410 2415	7485

5	GAA CA Glu Gl 2420			Arg					Glu					Lys	7533
10	AAC AT Asn Il		Gly					Asp					Ile		7581
10	AAT GA Asn Gl						Lys					Gln			7629
15	GTA AC		e Thr			Glu					Asp				7677
20	CTT CA Leu Gl 248	ln As:			Asp					Arg					7725
25	AGG CA Arg G 2500			Phe					Ser					Lys	7773
30	TCC AG		u Pro					Lys					Gly		7821
30	CCC TO						Gln					Gly			7869
35	CAT TO		e Lys			Ser					Ser				7917
40	ACT G. Thr G 25	lu As			Gly					Trp					7965
45	CAG T Gln L 2580			Gly					Pro					Lys	8013
50	GGA A Gly L		u Glu					Leu					Gly		8061
	CCA A Pro L			Ser			Trp					Tyr			8109
55	ATA T		s Leu			Met					Pro				8157
60	AAT A Asn A 26				Pro					Leu					8205

5					Ile					Arg					AAG Lys 2		8253
				Asp					Lys					Cys	GTT Val 2690		8301
10			Ile					Asn					Ser		AAT Asn		8349
15		Ser					Gln					Ile			ACA Thr		8397
20	Gly					Lys					Pro				GCT Ala		8445
25	Leu 2740	Lys	Asn	Gly	Arg 2	Leu 2745	Thr	Val	Gly	Gln 2	Lys 2750	Ile	Ile	Leu		Gly 2755	8493
	Ala	Glu	Leu	Val 2	Gly 2760	Ser	Pro	Asp	Ala 2	Cys 2765	Thr	Pro	Leu	Glu ?	GCC Ala 2770	Pro	8541
30			Leu					Ser					Arg		GCT Ala		8589
35		Tyr					Phe					Arg			CCT Pro		8637
40	Pro					Phe					Asn				GTT Val		8685
45	Val 2820	Ile	Ile	Gln	Arg	Ala 2825	Tyr	Pro	Ile	Gln 2	Trp 2830	Met	Glu	Lys		Ser 2835	8733
				Tyr					Glu					Lys	GAA Glu 2850		8781
50	Ala	Lys	Tyr	Val 2855	Glu	Ala	Gln	Gln	Lys 2860	Arg	Leu	Glu	Ala	Leu 2865	TTC Phe	Thr	8829
55		Ile					Glu					Asn			AAA Lys		8877
60			Pro			Ala					Gln				TTG Leu		8925
	GAT	GGT	GCA	GAG	CTT	TAT	GAA	GCA	GTG	AAG	AAT	GCA	GCA	GAC	CCA	GCT	8973

	Asp 2900	Gly	Ala	Glu		Tyr :905	Glu	Ala	Val		Asn 2910	Ala	Ala	Asp	Pro 2	Ala 915	
5				Gly					Glu					Leu	AAT Asn 2930		9021
10			Gln					Lys					Ile		TTG Leu	_	9069
15		Arg					Ser					Glu			TTA Leu		9117
20	Arg					Val					Ile				TCA Ser		9165
20					Ser					Ile					TCA Ser		9213
25				Leu					Lys					Tyr	CAT His 3010		9261
30			Ser					Lys					Asn		CAG Gln		9309
35		Ala					Gln					Pro			GAT Asp		9357
40	Ile					Tyr					Pro				AGC Ser		9405
40		Leu			Asp					Cys					CTA Leu		9453
45				Val					Lys					Pro	TTC Phe 3090		9501
50			Ser					Asn					Lys		TGG Trp		9549
55		Leu					Ile					Leu			GCA Ala		9597
60						Pro					Gly				TTA Leu		9645
															CAC His		9693

CTC AGA CTG AAA CGA CGT TGT ACT ACA TCT CTG ATC AAA GAA CAG GAG

Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys Glu Glu Glu

AGT TCC CAG GCC AGT ACG GAA GAA TGT GAG AAA AAT AAG CAG GAC ACA Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys Gln Asp Thr ATT ACA ACT AAA AAA TAT ATC TAA Ile Thr Thr Lys Lys Tyr Ile (2) INFORMATION FOR SEQ ID NO:13: (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 3418 amino acids (B) TYPE: amino acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear (ii) MOLECULE TYPE: protein (v) FRAGMENT TYPE: internal (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13: Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe Glu Glu Leu Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu Glu Ser Glu His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr Pro Gln Arg Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile Phe Lys Glu Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys Glu Leu Asp Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser Arg His Lys Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp Asp Val Ser Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val Val Leu Gln Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val Cys Gly Ser Leu Phe His Thr Pro Lys Phe Val Lys Gly Arg Gln Thr Pro Lys His Ile Ser Glu Ser Leu Gly Ala Glu Val Asp Pro Asp Met Ser Trp Ser Ser Ser Leu Ala Thr Pro Pro Thr Leu Ser Ser Thr Val Leu Ile Val Arg Asn Glu Glu Ala Ser Glu Thr Val Phe Pro His Asp Thr Thr Ala Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu Lys Lys Asn Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr Asn Gln Arg Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn Ser Phe Lys Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro His Val Leu Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu

	305	Asp				310					315					320
5		Lys			325					330					335	
J		Ala		340					345					350		
		Phe	355					360					365			
10		Val 370					375					380				
	385	Glu				390					395					400
15		Gly			405					410					415	
		Ser		420					425					430		
		Lys	435					440					445			
20		Ser 450					455					460				
	465	Asn				470					475					480
25		Leu			485					490					495	
		Phe		500					505					510		
	_	Glu	515					520					525			
30		Lys 530					535					540				
	545	Cys Ser				550					555					560
35		Ser			565					570					575	
		Ile		580					585					590		
40		Lys	595					600					605			
40		610 Glu					615					620				
	625					630					635					640
45		Ser			645					650					655	
				660					665					670		Tyr
50			675					680					685			Pro
		690					695	;				700				Asp
	705 Pro	Lys	Ser	Lys	Lys	710 Val		Asp	ıl∈	. Lys	715 Glu		. Val	. Leu	ı Ala	720 Ala
55	Ala	e Cys	His	Pro	725 Val		His	: Ser	Lys	730 Val		Tyr	Ser			Asp
	Ph€	e Gln	Ser	740 Gln		Ser	Leu				His	Glu				Thr
60	Leu				Pro	Thr				o Val	. Leu				ı Val	Met
	Ile	770 Ser		g Gly	Lys	Glu	775 1 Sei		Lys	Met	Ser	780 Asp		Leu	ı Lys	Gly

	785					790	3		_	-1	795		<b>-</b> 1 -	D	<b>3</b> 4 - +	800
			_		805					810					Met 815	
				820					825					830	Val	
			835					840					845		Arg	
10		850					855					860			Asn	
	865					870					875				Ser	880
					885					890					Ala 895	
15				900					905					910	Glu	
			915					920					925		Met	
20		930	_	_			935					940			Ile	
	945	-				950					955				Val	960
				_	965					970					Ile 975	
25				980					985					990	Asp	
	_		995					1000	)				100	5	Gly	
30		1010	)				1015	5				1020	)		Asn	
	1025	5				1030	)				1035	5			Pro	104
					1045	5				1050	)				Asn 1055	5
35	-	_		1060	)				1069	5				107		
			1075	5				108	0				108	5	Thr	
40			Leu	Phe	Ser	1.370	CID					Δcn	HIS	ASI	ьeu	THE
	_	109	)	_			109	5	Phe			110				
		Ser	)	Lys			1099 Ile	5				110 Thr			Glu	
	1109	Ser	) Gln	-	Ala	Glu 1110 Glu	109! Ile 0	5 Thr	Glu	Leu	Ser 111! Arg	110 Thr 5	Ile	Leu		Glu 112 Ile
45	1109 Ser	Ser Gly	Gln Ser	Gln Ser	Ala Phe 1129 Thr	Glu 1110 Glu	1099 Ile 0 Phe	Thr Thr	Glu Gln Pro	Leu Phe 113 Glu	Ser 111! Arg	110 Thr 5 Lys	Ile Pro	Leu Ser Thr	Glu Tyr 113! Ile	Glu 112 Ile
45	1109 Ser Leu	Ser Gly Gln	Gln Ser Lys Thr	Gln Ser 1140 Ser	Ala Phe 1129 Thr	Glu 1110 Glu 5 Phe	1099 Ile O Phe Glu	Thr Thr Val	Glu Gln Pro 1149 Asp	Leu Phe 113 Glu 5	Ser 111! Arg O Asn	1100 Thr 5 Lys Gln	Ile Pro Met His	Leu Ser Thr 115 Val	Glu Tyr 113! Ile 0	Glu 112 Ile 5
	1105 Ser Leu Lys	Ser Gly Gln Thr	Gln Ser Lys Thr 1159	Gln Ser 1140 Ser	Ala Phe 1129 Thr O Glu	Glu 1110 Glu 5 Phe Glu	1099 Ile O Phe Glu Cys	Thr Thr Val Arg 116 Val	Glu Gln Pro 114! Asp	Leu Phe 113 Glu 5 Ala	Ser 1119 Arg O Asn Asp	110 Thr Lys Gln Leu	Ile Pro Met His 116 Gln	Leu Ser Thr 115 Val	Glu Tyr 113! Ile O Ile	Glu 112 Ile 5 Leu
<b>4</b> 5	1105 Ser Leu Lys Asn Thr	Ser Gly Gln Thr Ala 117 Val	Gln Ser Lys Thr 1159 Pro	Gln Ser 1140 Ser Ser	Ala Phe 1129 Thr Glu Ile	Glu 1110 Glu 5 Phe Glu Gly Arg	1099 Ile O Phe Glu Cys Gln 117 Lys	Thr Thr Val Arg 116 Val	Glu Gln Pro 1149 Asp O Asp	Leu Phe 113 Glu 5 Ala Ser	Ser 111! Arg O Asn Asp Ser Leu	Thr  Lys  Gln  Leu  Lys  118  Leu	Ile Pro Met His 116 Gln	Leu Ser Thr 115 Val 5 Phe	Glu Tyr 113: Ile 0 Ile Glu	Glu 112 Ile 5 Leu Met Gly Cys
	Leu Lys Asn Thr	Ser Gly Gln Thr Ala 117 Val	Gln Ser Lys Thr 1159 Pro Glu	Gln Ser 1140 Ser Ser The ser	Ala Phe 1129 Thr Glu Ile Lys Ser	Glu 1110 Glu 5 Phe Glu Gly Arg 119 Gly	1099 Ile O Phe Glu Cys Gln 117 Lys	Thr Thr Val Arg 116 Val 5	Glu Gln Pro 114! Asp O Asp	Phe 113 Glu 5 Ala Ser Gly Asp	Ser 111! Arg O Asn Asp Ser Leu 119 Glu	1100 Thr 5 Lys Gln Leu Lys 118 Leu 5	Ile Pro Met His 116 Gln O Lys	Leu Ser Thr 115 Val 5 Phe Asn	Glu Tyr 113: Ile 0 Ile Glu Asp	Glu 112 Ile 5 Leu Met Gly Cys 120 Phe
	Leu Lys Asn Thr 1189	Ser Gly Gln Thr Ala 117 Val 5	Gln Ser Lys Thr 115: Pro Glu Ser	Gln Ser 1140 Ser 5 Ser Ile Ala Tyr	Ala Phe 1129 Thr Glu Ile Lys Ser 1200 Ser	Glu 1110 Glu 5 Phe Glu Gly Arg 119 Gly	1099 Ile O Phe Glu Cys Gln 117 Lys O	Thr Thr Val Arg 116 Val Fhe	Glu Gln Pro 114! Asp O Asp Ala	Phe 113 Glu 5 Ala Ser Gly Asp 121 Lys	Ser 111: Arg O Asn Asp Ser Leu 119 Glu	1100 Thr 5 Lys Gln Leu Lys 118 Leu 5 Asn	Ile Pro Met His 116 Gln O Lys Glu	Leu Ser Thr 115 Val 5 Phe Asn Val	Glu Tyr 113: Ile 0 Ile Glu Asp Gly 121 Thr	Glu 112 Ile 5 Leu Met Gly Cys 120 Phe
50	Leu Lys Asn Thr 1189 Asn	Ser Gly Gln Thr Ala 117 Val 5 Lys	Gln Ser Lys Thr 115: Pro Glu Ser Phe	Gln Ser 1140 Ser Ser Ile Ala Tyr 122 Lys	Ala Phe 1129 Thr Glu Ile Lys Ser 1209 Ser	Glu 1110 Glu 5 Phe Glu Gly Arg 119 Gly 5	1099 Ile O Phe Glu Cys Gln 117 Lys O Tyr His	Thr Thr Val Arg 116 Val Phe Leu Gly	Glu Gln Pro 1149 Asp O Asp Ala Thr Thr 122 Phe	Leu Phe 113 Glu  Ala Ser Gly Asp 121 Lys 5	Ser 111: Arg O Asn Asp Ser Leu 119 Glu O Leu	1100 Thr 5 Lys Gln Leu Lys 118 Leu 5 Asn	Ile Pro Met His 116 Gln O Lys Glu Val	Leu Ser Thr 115 Val 5 Phe Asn Val Ser 123 Asn	Glu Tyr 113: Ile 0 Ile Glu Asp Gly 121 Thr	Glu 112 Ile 5 Leu Met Gly Cys 120 Phe
50	Leu Lys Asn Thr 1189 Asn Arg Ala	Ser Gly Gln Thr Ala 117 Val 5 Lys Gly Leu Glu 125	Gln Ser Lys Thr 1159 Pro Glu Ser Phe Gln 123 Thr	Gln Ser 1140 Ser Ser Ile Ala Tyr 122 Lys Ser	Ala Phe 1129 Thr Glu Ile Lys Ser 1200 Ser Ala Ala	Glu 1110 Glu 5 Phe Glu Gly Arg 119 Gly 5 Ala Val	1099 Ile O Phe Glu Cys Gln 117 Lys O Tyr His Lys Val 125	Thr Thr Val Arg 116 Val Fhe Leu Gly Leu 124 His	Glu Gln Pro 1149 Asp O Asp Ala Thr Thr 122 Phe O Pro	Leu Phe 113 Glu  Ala Ser Gly Asp 121 Lys Ser Ile	Ser 111: Arg Asn Asp Ser Leu 119 Glu O Leu Asp	1100 Thr Lys Gln Leu Lys 118 Leu Asn Ale Leu 126	Ile Pro Met His 116 Gln O Lys Glu Val Glu 124 Ser	Leu Ser Thr 115 Val 5 Phe Asn Val Ser 123 Asn 5 Ser	Glu Tyr 113: Ile 0 Ile Glu Asp Gly 121 Thr 0 Ile Ser	Glu 112 Ile 5 Leu Met Gly Cys 120 Phe 5 Glu Ser Lys
50 55	Leu Lys Asn Thr 1189 Asn Arg Ala	Ser Gly Gln Thr Ala 117 Val 5 Lys Gly Leu Glu 125 His	Gln Ser Lys Thr 1159 Pro Glu Ser Phe Gln 123 Thr	Gln Ser 1140 Ser Ser Ile Ala Tyr 122 Lys Ser	Ala Phe 1129 Thr Glu Ile Lys Ser 1200 Ser Ala Ala	Glu 1110 Glu 5 Phe Glu Gly Arg 119 Gly 5 Ala Val	1099 Ile O Phe Glu Cys Gln 117 Lys O Tyr His Lys Val 125 Ser	Thr Thr Val Arg 116 Val Fhe Leu Gly Leu 124 His	Glu Gln Pro 1149 Asp O Asp Ala Thr Thr 122 Phe O Pro	Leu Phe 113 Glu  Ala Ser Gly Asp 121 Lys Ser Ile	Ser 111: Arg Asn Asp Ser Leu 119 Glu O Leu Asp	Thr Lys Gln Leu Lys 118 Leu Asn Asn Ile Leu 126 Glu	Ile Pro Met His 116 Gln O Lys Glu Val Glu 124 Ser	Leu Ser Thr 115 Val 5 Phe Asn Val Ser 123 Asn 5 Ser	Glu Tyr 113: Ile 0 Ile Glu Asp Gly 121 Thr 0 Ile Ser	Glu 112 Ile 5 Leu Met Gly Cys 120 Phe 5 Glu Ser

1285 1290  Asn Ile Glu Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr 5 1300 1305 131  Tyr Lys Arg Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala	1295 Glu Asn
	0
1315 1320 1325	
Arg Asn Ser His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser 1330 1335 1340	
Asp Thr Val Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe 1345 1350 1355	136
Gln His Asn Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys 1365 1370	1375
Asn Thr Gln Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu 15 1380 1385 139	0
Ala Lys Ala Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys 1395 1400 1405	
Leu Thr Ala Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu 1410 1415 1420  20 Asp Thr Phe Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val	
20 Asp Thr Phe Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val 1425 1430 1435 Glu Ser Phe Asn Lys Ile Val Asn Phe Phe Asp Gln Lys Pro	144
1445 1450  Leu His Asn Phe Ser Leu Asn Ser Glu Leu His Ser Asp Ile	1455
25 1460 1465 147 Asn Lys Met Asp Ile Leu Ser Tyr Glu Glu Thr Asp Ile Val	0
1475 1480 1485 Lys Ile Leu Lys Glu Ser Val Pro Val Gly Thr Gly Asn Gln	
1490 1495 1500	
30 Thr Phe Gln Gly Gln Pro Glu Arg Asp Glu Lys Ile Lys Glu 1505 1510 1515  Leu Leu Gly Phe His Thr Ala Ser Gly Lys Lys Val Lys Ile	152
1525 1530 Glu Ser Leu Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu	1535
35 1540 1545 155  Thr Ser Glu Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr	0
1555 1560 1565  Tyr Arg Glu Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr	
1570 1575 1580 40 Ile Thr Ala Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu	
1585 1590 1595 Asp Lys Asn Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys	160
1605 1610 Ser Asp Asn Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser	1615
45 1620 1625 163  Ile Phe Leu Lys Val Lys Val His Glu Asn Val Glu Lys Glu	0
1635 1640 1645 Lys Ser Pro Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser	
1650 1655 1660 50 Glu Asn Ser Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys	Thr Ser
1665 1670 1675 Val Ser Gln Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg	168 Glu Gly
1685 1690 Ile Phe Asp Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr	1695 Val Gly
55 1700 1705 171 Asn Tyr Leu Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu	<del>-</del>
1715 1720 1725  Lys Asn His Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asr	ı Ser Ser
1730 1735 1740	Asp Ser
60 Met Ser Asn Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asr 1745 1750 1755	176

		1765		1	770		1775
	-	Glu Asp G 1780		1785	er Phe Ser	1790	
5	179	5	180	00		1805	
	1810		1815		Ser Pro Cys 1820		
10	1825	1	830		Ser Asn Asn 1835		184
		1845		1	ys Ile Val .850		1855
a ==		1860		1865	The Thr Asp	1870	ı
15	187	5	188	80	Ser Lys Ile	1885	
	1890		1895		1900 Ser Thr His	l	
20	1905	1	.910		1915		192
		1925		1	Leu Gln His		1935
	_	1940		1945	Ser Pro Cys	1950	)
25	195	55	19	60	le Gly Lys	1965	
	1970		1975		he Ser Thr 1980	)	
30	Ser Val Glr 1985		Asp Ala Se 1990	r Leu G	In Asn Ala 1995	arg Gin	200
	Ser Glu Ile	2005		2	Val Phe Ser 2010		2015
		2020		2025	Thr Arg Glu	2030	)
35	203	35	20	40	Gln Lys Gly	2045	
	Val Val Ası 2050	ı Ser Ser A	Ala Phe Se 2055	r Gly I	Phe Ser Thr 2060		GIA TAR
40	2065	2	2070		His Lys Val 2075		208
		2085	_	2	His Ser Leu 2090		2095
		2100		2105	Leu Pro Arg	2110	כ
45	Asn Pro Gli			r Glu N .20	Met Glu Lys	2125	ser Lys
	2130		2135		Val Glu Gly 2140	)	
50	Asn Asn Hi 2145		Lys Val Se 2150	r Pro T	Tyr Leu Ser 2155	GIn Phe	Gin Gin 216
	Asp Lys Gl	2165		2	Lys Val Ser 2170		2175
		2180		2185	Ser Pro Lys	219	0
55	21	95	22	00	Asp Val Pro	2205	
	2210		2215		Asp Ser Glu 222 Phe Met Glu	0	
60	2225	2	2230		2235		224
	Thr Asp Se	r Lys Leu I 2245	Pro Ser Hi		Thr His Ser 2250	Leu Phe	Thr Cys 2255

	Pro	Glu	Asn	Glu 2260		Met	Val	Leu	Ser 2265		Ser	Arg	Ile	Gly 2270		Arg
5	Arg	Gly	Glu 2275	Pro		Ile		Val 2280	Gly		Pro	Ser	Ile 2285		Arg	Asn
J	Leu	Leu 2290	Asn	Glu	Phe	Asp		Ile		Glu	Asn	Gln 2300	Glu		Ser	Leu
	Lys 2305	Ala		Lys	Ser	Thr 2310	Pro		Gly	Thr	Ile 2315		Asp	Arg	Arg	Leu 232
10			His	His	Val 2325	Ser		Glu	Pro	Ile 2330		Cys	Val	Pro	Phe 2335	
	Thr	Thr	Lys	Glu 2340		Gln	Glu	Ile	Gln 2345		Pro	Asn	Phe	Thr 2350		Pro
15	Gly	Gln	Glu 2355	Phe		Ser	Lys	Ser 2360		Leu	Tyr	Glu	His 2365		Thr	Leu
	Glu	Lys 2370	Ser	Ser	Ser	Asn	Leu 2375		Val	Ser	Gly	His 2380		Phe	Tyr	Gln
	Val 2389		Ala	Thr	Arg	Asn 2390		Lys	Met	Arg	His 2399		Ile	Thr	Thr	Gly 240
20	Arg	Pro	Thr	Lys	Val 2405		Val	Pro	Pro	Phe 2410		Thr	Lys	Ser	His 2415	
	His	Arg	Val	Glu 2420		Cys	Val	Arg	Asn 2425		Asn	Leu	Glu	Glu 2430		Arg
25	Gln	Lys	Gln 2435	Asn	Ile	Asp	Gly	His 2440		Ser	Asp	Asp	Ser 2445		Asn	Lys
	Ile	Asn 2450		Asn	Glu	Ile	His 2455		Phe	Asn		Asn 2460		Ser	Asn	Gln
	Ala 246		Ala	Val	Thr	Phe 2470		Lys	Cys	Glu	Glu 2475		Pro	Leu	Asp	Leu 248
30			Ser	Leu	Gln 2485	Asn		Arg	Asp	Ile 2490		Asp	Met	Arg	Ile 2495	
	Lys	Lys	Gln	Arg 2500		Arg	Val	Phe	Pro 250		Pro	Gly	Ser	Leu 2510		Leu
35		•	251					2520	)				252	5		
	_	2530	)	Pro			2535	5				2540	)			
	254	5		His		2550	)				255	5				256
40				Thr	2569	5				2570	)				2575	5
				Gln 2580	)				258	5				2590	0	
45	_	-	259		_			260	)				260	5		
	•	261	ס -	Pro	•		261	5	_		_	262	0			
	262	5		Ile	_	263	0				263	5				264
50				Asn -	264	5				2650	)				265	5
		_	_	Tyr 2660	C				266	5				267	0	
55	_	_	267			_		268	0				268	5		
	-	269	0	Asp			269	5				270	0			
	270	5		Thr		271	0				271	5				272
60				Gly	272	5				273	0				273	5
	Leu	Ala	Val	Leu	Lys	Asn	Gly	Arg	Leu	Thr	Val	Gly	Gln	Lys	Ile	Ile

				2740					2745	;				2750		
	Leu	His	Gly 2755	Ala	Glu	Leu		Gly 2760		Pro	Asp	Ala	Cys 2765		Pro	Leu
5	Glu	Ala 2770		Glu	Ser	Leu	Met 2775		Lys	Ile	Ser	Ala 2780		Ser	Thr	Arg
	2785	5	_	Trp	_	2790	)				2795					280
10	Phe	Pro	Leu	Pro	Leu 2805		Ser	Leu	Phe	Ser 2810		Gly	Gly	Asn	Val 2815	
	_		_	Val 2820	1				2825	5				2830		
			2835					2840	)				2845	5		
15	-	2850	)	Ala			2855	5				2860	)			
	286	5		Lys		2870	)				2875	5				288
20		_		Tyr	2885	5				2890	)				2895	5
				Asp 2900	)				2905	5				2910	)	
0.5	_		2915					2920	)				2925	5		
25		2930	)	His			2935	5				2940	)			
	294	5		Ile		295	0				2955	5				296
30	_			Arg Lys	2969	5				2970	)				297	5
				2980 Leu	)				298	5				2990	)	
35			299	5				3000	)				3005	5		Asn
33	_	301	0				301	5				302	0			Val
	302	5		Ile		303	0				303	5				304
40		_		Phe	304	5				3050	)				305	5
				3060 Gly	<b>O</b>				306	5				3070	)	
45			307	5				308	0				308	5		Lys
	Phe	309 Trp		Asp	Leu	Asn	309 Glu		Ile	Ile	Lys	310 Pro		Met	Leu	Ile
	310 Ala		Ser	Asn	Leu	311 Gln		Arg	Pro	Glu	311 Ser		Ser	Gly	Leu	312 Leu
50	Thr	Leu	Phe	Ala	312 Gly		Phe	Ser	Val	313 Phe		Ala	Ser	Pro	313 Lys	5 Glu
	Gly	His	Phe	314 Gln		Thr	Phe	Asn	314 Lys		Lys	Asn				Asn
55	Ile				Cys	Asn				Asn	Lys				Ile	Leu
				Asp	Pro				Thr	Pro				Cys	Thr	Ser
60	318 Gly		Tyr	Thr	Ala 320			Ile	Pro	Gly 321			Asn	Lys	Leu 321	320 Leu
60	Met	Ser	Ser	Pro	Asn		Glu	Ile	Tyr 322	Tyr		Ser	Pro	Leu 323	Ser	Leu

	3235 3240 3245	ACC
5	Thr Ser Lys Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys 3250 3255 3260	Asn
	Cys Lys Lys Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu	Pro
		328
	Pro Pro Val Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln 3285 3290 3295	Lys
10	Ala Phe Gln Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro 3300 3305 3310	Ile
	Lys Lys Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys 3315 3320 3325	Phe
15	Asn Glu Ile Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu 3330 3335 3340	Leu
	Ala Leu Ile Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu	Lys
		336
	Gln Phe Ile Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser 3365 3370 3375	
20	Glu Asp Tyr Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile 3380 3385 3390	
	Glu Gln Glu Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn	Lys
	3395 3400 3405	
25	Gln Asp Thr Ile Thr Thr Lys Lys Tyr Ile 3410 3415	
	(2) INFORMATION FOR SEQ ID NO:14:	
	(:) CROUDINGE CHARACTERICTICS.	
30	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li></ul>	
50	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: single	
	(D) TOPOLOGY: linear	
35		
33	(A) NAME/KEY:	
	(B) LOCATION:	
	(D) OTHER INFORMATION: 2F primer	
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:	
	TGAGTTTTAC CTCAGTCACA	20
4 =	(2) INFORMATION FOR SEQ ID NO:16:	
45	(i) SEQUENCE CHARACTERISTICS:	
	(A) LENGTH: 41 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: single	
50	(D) TOPOLOGY: linear	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:	
55	CAGGAAACAG CTATGACCCT GTGACGTACT GGGTTTTTAG C	41
	(2) INFORMATION FOR SEQ ID NO:17:	
	(i) SEQUENCE CHARACTERISTICS:	
60	(A) LENGTH: 24 base pairs	
	(B) TYPE: nucleic acid	
	(C) STRANDEDNESS: single	

5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 3FII primer</li></ul>	
7.0	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:	
10	GATCTTTAAC TGTTCTGGGT CACA	24
	(2) INFORMATION FOR SEQ ID NO:18:	
15	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 22 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single	
20	(D) TOPOLOGY: linear	
٥٢	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 3RII primer</li></ul>	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:	
	CCCAGCATGA CACAATTAAT GA	22
30	(2) INFORMATION FOR SEQ ID NO:19:	
35	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 44 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
40	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 4F/M 13F primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:	
45	TGTAAAACGA CGGCCAGTAG AATGCAAATT TATAATCCAG AGTA	44
	(2) INFORMATION FOR SEQ ID NO:20:	
50	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
55	/a \ Name /kek	
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 4R-1A primer</li></ul>	
60	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:	
	ATCAGATTCA TCTTTATAGA AC	22

(D) TOPOLOGY: linear

	(2) INFORMATION FOR SEQ ID NO:21:	
5	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 40 base pairs	
	<ul><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
10	(D) TOPOLOGI: Timeat	
15	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 5+6F/M13F primer</li></ul>	
13	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:	
	TGTAAAACGA CGGCCAGTTG TGTTGGCATT TTAAACATCA	40
20	(2) INFORMATION FOR SEQ ID NO:22:	
25	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 38 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
30	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 5+6R/M13R primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:	
35	CAGGAAACAG CTATGACCCA GGGCAAAGGT ATAACGCT	38
	(2) INFORMATION FOR SEQ ID NO:23:	
40	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 38 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
45	. (A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 7F/M13F primer	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:	
	TGTAAAACGA CGGCCAGTTA AGTGAAATAA AGAGTGAA	38
55	(2) INFORMATION FOR SEQ ID NO:24:	
	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 36 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
60	(D) TOPOLOGY: linear	

5	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 7R/M13R primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:  CAGGAAACAG CTATGACCAG AAGTATTAGA GATGAC	36
10	(2) INFORMATION FOR SEQ ID NO:25:	
10		
15	<ul> <li>(i) SEQUENCE CHARACTERISTICS:</li> <li>(A) LENGTH: 40 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul>	
20	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 8F/M13F primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:	
25	TGTAAAACGA CGGCCAGTGC CATATCTTAC CACCTTGTGA	40
	(2) INFORMATION FOR SEQ ID NO:26:	
30	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
35	(ix) FEATURE:	
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 8FIA primer</li></ul>	
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:	
	TTGCATTCTA GTGATAATAT AC	22
45	(2) INFORMATION FOR SEQ ID NO:27:	
13	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 19 base pairs	
50	(B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
55	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 8RIA primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:	
60	AATTGTTAGC AATTTCAAC	19
	(2) INFORMATION FOR SEQ ID NO:28:	

5	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 40 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
10	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 9F/M13F primer</li></ul>	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28: TGTAAAACGA CGGCCAGTTG GACCTAGGTT GATTGCAGAT	40
20	<ul> <li>(2) INFORMATION FOR SEQ ID NO:29:</li> <li>(i) SEQUENCE CHARACTERISTICS: <ul> <li>(A) LENGTH: 40 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul> </li> </ul>	
23	(A) NAME/KEY:	
30	<ul><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 9R/M13R primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:</li></ul>	
	CAGGAAACAG CTATGACCTA AACTGAGATC ACGGGTGACA	40
35	(2) INFORMATION FOR SEQ ID NO:30:	
40	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 24 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
45	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 10AF primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:	
50	GAATAATATA AATTATATGG CTTA	24
	(2) INFORMATION FOR SEQ ID NO:31:	
55	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 37 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
60	(A) NAME/KEY:	
	(B) LOCATION:	

	(D) OTHER INFORMATION: 10AR/M13R primer	
5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:	
5	CAGGAAACAG CTATGACCCC TAGTCTTGCT AGTTCTT	37
	(2) INFORMATION FOR SEQ ID NO:32:	
10	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 42 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
15		
20	<pre>(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 10BF/M13F primer (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:</pre>	
		42
	TGTAAAACGA CGGCCAGTAR CTGAAGTGGA ACCAAATGAT AC	42
25	(2) INFORMATION FOR SEQ ID NO:33:	
30	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 44 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
35	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 10BR/M13R primer</li></ul>	
35	(B) LOCATION: (D) OTHER INFORMATION: 10BR/M13R primer  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:	
35 40	(B) LOCATION: (D) OTHER INFORMATION: 10BR/M13R primer	44
	(B) LOCATION: (D) OTHER INFORMATION: 10BR/M13R primer  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:	44
	(B) LOCATION: (D) OTHER INFORMATION: 10BR/M13R primer  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:  CAGGAAACAG CTATGACCAC GTGGCAAAGA ATTCTCTGAA GTAA	44
40	(B) LOCATION: (D) OTHER INFORMATION: 10BR/M13R primer  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:  CAGGAAACAG CTATGACCAC GTGGCAAAGA ATTCTCTGAA GTAA  (2) INFORMATION FOR SEQ ID NO:34:  (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 40 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single	44
40 45 50	(B) LOCATION: (D) OTHER INFORMATION: 10BR/M13R primer  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:  CAGGAAACAG CTATGACCAC GTGGCAAAGA ATTCTCTGAA GTAA  (2) INFORMATION FOR SEQ ID NO:34:  (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 40 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	44
40 45	(B) LOCATION: (D) OTHER INFORMATION: 10BR/M13R primer  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:  CAGGAAACAG CTATGACCAC GTGGCAAAGA ATTCTCTGAA GTAA  (2) INFORMATION FOR SEQ ID NO:34:  (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 40 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear  (ix) FEATURE:  (A) NAME/KEY: (B) LOCATION:	44
40 45 50	(B) LOCATION: (D) OTHER INFORMATION: 10BR/M13R primer  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:  CAGGAAACAG CTATGACCAC GTGGCAAAGA ATTCTCTGAA GTAA  (2) INFORMATION FOR SEQ ID NO:34:  (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 40 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear  (ix) FEATURE:  (A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 10CF/M13F primer	44
40 45 50	(B) LOCATION: (D) OTHER INFORMATION: 10BR/M13R primer  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:  CAGGAAACAG CTATGACCAC GTGGCAAAGA ATTCTCTGAA GTAA  (2) INFORMATION FOR SEQ ID NO:34:  (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 40 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear  (ix) FEATURE:  (A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 10CF/M13F primer  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:	

(A) LENGTH: 19 base pairs

	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:	
5	AAGAAGCAAA ATGTAATAAG GA	22
כ	(2) INFORMATION FOR SEQ ID NO:39:	
10	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
15	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11BR primer</li></ul>	
20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:	
20	CATTTAAAGC ACATACATCT TG	22
	(2) INFORMATION FOR SEQ ID NO:40:	
25	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
30	(D) TOPOLOGY: linear	
35	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11CF primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:</li></ul>	
	TCTAGAGGCA AAGAATCATA C	21
40	(2) INFORMATION FOR SEQ ID NO:41:	
45	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
50	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 11CR primer	
cc	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:	22
55	CAAGATTATT CCTTTCATTA GC	22
60	<ul> <li>(2) INFORMATION FOR SEQ ID NO:42:</li> <li>(i) SEQUENCE CHARACTERISTICS: <ul> <li>(A) LENGTH: 22 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> </ul> </li> </ul>	

5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11DF primer</li></ul>	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:	2.2
	AACCAAAACA CAAATCTAAG AG	22
	(2) INFORMATION FOR SEQ ID NO:43:	
15	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 23 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
20	(D) TOPOLOGY: linear	
20	(A) NAME/KEY: (B) LOCATION:	
	(D) OTHER INFORMATION: 11DR primer	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:	
	GTCATTTTTA TATGCTGCTT TAC	23
30	(2) INFORMATION FOR SEQ ID NO:44:	
35	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
40	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11EF primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:	
45	GGTTTTATAT GGAGACACAG G	21
	(2) INFORMATION FOR SEQ ID NO:45:	
50	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 23 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
55		
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11ER primer</li></ul>	
60	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:	
		23

(D) TOPOLOGY: linear

	(2) INFORMATION FOR SEQ ID NO:46:	
5	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
10		
15	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11FF primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:</li></ul>	
	(XI) DECORNCE DESCRIPTION. ORG. 15 No. 40.	
	ATCACAGTTT TGGAGGTAGC	20
20	(2) INFORMATION FOR SEQ ID NO:47:	
25	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
30	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11FR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:	
35	CTGACTTCCT GATTCTTCTA A	21
	(2) INFORMATION FOR SEQ ID NO:48:	
40	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
45		
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11GF primer</li></ul>	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:	
	CTCAGATGTT ATTTTCCAAG C	21
55	<ul><li>(2) INFORMATION FOR SEQ ID NO:49:</li><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li></ul>	
	(B) TYPE: nucleic acid (C) STRANDEDNESS: single	
60	(D) TODOLOGY: linear	

5	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 11GR primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:	21
	CTGTTAAATA ACCAGAAGCA C	21
10	(2) INFORMATION FOR SEQ ID NO:50:	
15	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 18 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
20	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11HF primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:	
25	AGGTAGACAG CAGCAAGC	18
	(2) INFORMATION FOR SEQ ID NO:51:	
30	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
35	(ix) FEATURE:	
	<ul><li>(A) NAME/KEY: None</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11HR primer</li></ul>	
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:	
	GTAATATCAG TTGGCATTTA TT	22
45	(2) INFORMATION FOR SEQ ID NO:52:	
50	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
55	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 111F primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:	
60	TGCAGAGGTA CATCCAATAA G	21
	(2) INFORMATION FOR SEQ ID NO:53:	

5	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
10	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11IR primer</li></ul>	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:  GATCAGTAAA TAGCAAGTCC G  (2) INFORMATION FOR SEQ ID NO:54:	21
20	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 23 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
25	(D) TOPOLOGY: linear	
30	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11JF primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:</li></ul>	
	TACTGAAAAT GAAGATAACA AAT	23
35	(2) INFORMATION FOR SEQ ID NO:55:	
40	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
45	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: !!JR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:	
50	ATTTTGTTCT TTCTTATGTC AG	22
55	<ul> <li>(2) INFORMATION FOR SEQ ID NO:56:</li> <li>(i) SEQUENCE CHARACTERISTICS: <ul> <li>(A) LENGTH: 35 base pairs</li> <li>(B) TYPE: nucleic acid</li> <li>(C) STRANDEDNESS: single</li> <li>(D) TOPOLOGY: linear</li> </ul> </li> </ul>	
60	(A) NAME/KEY:	
	(B) LOCATION:	

	(D) OTHER INFORMATION: 11KF-M13 primer	
5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:	
5	TGTAAAACGA CGGCCAGTCT ACTAAAACGG AGCAA	35
	(2) INFORMATION FOR SEQ ID NO:57:	
10	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 35 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
15		
20	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11KR-M13 primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:</li></ul>	
	CAGGAAACAG CTATGACCGT ATGAAAACCC AACAG	35
25	(2) INFORMATION FOR SEQ ID NO:58:	
30	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
35	(A) NAME/KEY: (B) LOCATION:	
	(D) OTHER INFORMATION: 11LF primer	
40	(D) OTHER INFORMATION: 11LF primer	22
40	(D) OTHER INFORMATION: 11LF primer  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:	22
40 45	(D) OTHER INFORMATION: 11LF primer  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:  CACAAAATAC TGAAAGAAAG TG	22
	(D) OTHER INFORMATION: 11LF primer  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:  CACAAAATAC TGAAAGAAAG TG  (2) INFORMATION FOR SEQ ID NO:59:  (i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 19 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear	22
45	(D) OTHER INFORMATION: 11LF primer  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:  CACAAAATAC TGAAAGAAAG TG  (2) INFORMATION FOR SEQ ID NO:59:  (i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 19 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single	22
45	(D) OTHER INFORMATION: 11LF primer  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:  CACAAAATAC TGAAAGAAAG TG  (2) INFORMATION FOR SEQ ID NO:59:  (i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 19 base pairs  (B) TYPE: nucleic acid  (C) STRANDEDNESS: single  (D) TOPOLOGY: linear  (A) NAME/KEY:  (B) LOCATION:	22
45	(D) OTHER INFORMATION: 11LF primer  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:  CACAAAATAC TGAAAGAAAG TG  (2) INFORMATION FOR SEQ ID NO:59:  (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 19 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear  (A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 11LR primer	22
45	(D) OTHER INFORMATION: 11LF primer  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:  CACAAAATAC TGAAAGAAAG TG  (2) INFORMATION FOR SEQ ID NO:59:  (i) SEQUENCE CHARACTERISTICS: (A) LENGTH: 19 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear  (A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 11LR primer  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:	

5	<ul><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
5		
10	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11MF primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:	
	GCAAAGACCC TAAAGTACAG	20
15	(2) INFORMATION FOR SEQ ID NO:61:	
20	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
25	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11MR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:	
2.0	CLEGALLER MOCEENCE LG	
30	CATCAAATAT TCCTTCTCTA AG	22
	(2) INFORMATION FOR SEQ ID NO:62:	
35	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 35 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
40		
40	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11NF-M13 primer</li></ul>	
45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:	
	TGTAAAACGA CGGCCAGTGA AAATTCAGCC TTAGC	35
	(2) INFORMATION FOR SEQ ID NO:63:	
50	(i) CEQUENCE CHARACTERICATION.	
	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 35 base pairs</li></ul>	
	(B) TYPE: nucleic acid	
55	<ul><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
	(A) NAME/KEY:	
60	(B) LOCATION:	
60	(D) OTHER INFORMATION: 11NR-M13 primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:	

35

CAGGAAACAG CTATGACCAT CAGAATGGTA GGAAT

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11PR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:	
10	GTCAGCAAAA ACCTTATGTG	20
	(2) INFORMATION FOR SEQ ID NO:68:	
15	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
20		
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11QF primer</li></ul>	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:	
	ACGAAAATTA TGGCAGGTTG T	21
30	(2) INFORMATION FOR SEQ ID NO:69:	
30	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	
35	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
40	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11QR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:	
45	CTTGTCTTGC GTTTTGTAAT G	21
	(2) INFORMATION FOR SEQ ID NO:70:	
50	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
55	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11RF primer</li></ul>	
60	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:	
50	GCTTCATAAG TCAGTCTCAT	20

	(2) INFORMATION FOR SEQ ID NO:71:	
5	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
10	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11RR primer</li></ul>	
15	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:	
	TCAAATTCCT CTAACACTCC	20
20	<ul><li>(2) INFORMATION FOR SEQ ID NO:72:</li><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 35 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	
25	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
30	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11SF-M13 primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:</li></ul>	
35	TGTAAAACGA CGGCCAGTTA CAGCAAGTGG AAAGC (2) INFORMATION FOR SEQ ID NO:73:	35
40	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 37 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
45	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11SR-M13 primer</li></ul>	
50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:  CAGGAAACAG CTATGACCAA GTTTCAGTTT TACCAAT  (2) INFORMATION FOR SEQ ID NO:74:	37
55	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 22 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
60	(D) TOPOLOGY: linear	

(A) NAME/KEY:

	(B) LOCATION: (D) OTHER INFORMATION: 11TF primer	
5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:	
	GTTCTTCAGA AAATAATCAC TC	22
	(2) INFORMATION FOR SEQ ID NO:75:	
10	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	
15	<ul><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
20	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11TR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:	
25	TGTAAAAAGA GAATGTGTGG C	21
23	(2) INFORMATION FOR SEQ ID NO:76:	
30	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 39 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
35	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11UF-M13 primer</li></ul>	
4.0	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:	
40	TGTAAAACGA CGGCCAGTAC TTTTTCTGAT GTTCCTGTG	39
	(2) INFORMATION FOR SEQ ID NO:77:	
45	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 39 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
50	(D) TOPOLOGY: linear	
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 11UR-M13 primer</li></ul>	
55	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:	
	CAGGAAACAG CTATGACCTA AAAATAGTGA TTGGCAACA	39
60	(2) INFORMATION FOR SEQ ID NO:78:	
	(i) SEQUENCE CHARACTERISTICS:	

	5	(A) LENGTH: 42 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
	10	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 12F/M13F primer</li></ul>	
		(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:	
		TGTAAAACGA CGGCCAGTAG TGGTGTTTTA AAGTGGTCAA AA	42
	15	(2) INFORMATION FOR SEQ ID NO:79:	
	20	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 40 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
deal was specifically	25	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 12R/M13R primer</li></ul>	
:	2.0	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:	
	30	CAGGAAACAG CTATGACCGG ATCCACCTGA GGTCAGAATA	40
		(2) INFORMATION FOR SEQ ID NO:80:	
	35	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 21 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
	40	(D) TOPOLOGY: linear	
		<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 13-2F primer</li></ul>	
	45	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:	
		TAACATTTAA GCATCCGTTA C	21
	50	(2) INFORMATION FOR SEQ ID NO:81:	
	55	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 28 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
	60	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 13-2R primer</li></ul>	

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 15FUT/M13-R primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:	
10	CAGGAAACAG CTATGACCAC TCTGTCATAA AAGCCATC	38
	(2) INFORMATION FOR SEQ ID NO:86:	
15	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 24 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
20	(D) TOPOBOGI. TIMEMI	
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 16AF primer</li></ul>	
25	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:	
	TTTGGTTTGT TATAATTGTT TTTA	24
30	(2) INFORMATION FOR SEQ ID NO:87:	
	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	
35	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
40	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 16AR primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:	
45	CCAACTTTTT AGTTCGAGAG	20
	(2) INFORMATION FOR SEQ ID NO:88:	
50	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 19 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
55	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 17F primer</li></ul>	
60	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:	
	TTCAGTATCA TCCTATGTG	19

(D) TOPOLOGY: linear

		(2) INFORMATION FOR SEQ ID NO:89:	
	5	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
	10	(D) TOPOLOGY: linear	
	15	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 17AR primer</li></ul>	
		(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:	
		AGAAACCTTA ACCCATACTG	20
	20	(2) INFORMATION FOR SEQ ID NO:90:	
	25	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 39 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
	30	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 18FUT/M13-AF primer</li></ul>	
<b>-</b>		(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:	
Varde Grade Saran Saran made Bardh	35	TGTAAAACGA CGGCCAGTGA ATTCTAGAGT CACACTTCC	39
		(2) INFORMATION FOR SEQ ID NO:91:	
Sud-Audi	40	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 38 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
	45		
		<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 18R/M13R primer</li></ul>	
	50	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:	
		CAGGAAACAG CTATGACCTT TAACTGAATC AATGACTG	38
	E E	(2) INFORMATION FOR SEQ ID NO:92:	
	55	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 41 base pairs  (B) TYPE: nucleic acid	
	60	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	

5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 19F/M13F primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:</li></ul>	
	TGTAAAACGA CGGCCAGTAA GTGAATATTT TTAAGGCAGT T	41
10	(2) INFORMATION FOR SEQ ID NO:93:	
15	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 39 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
20	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 19FUT/M13-R primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:	
25	CAGGAAACAG CTATGACCAA GAGACCGAAA CTCCATCTC	39
	(2) INFORMATION FOR SEQ ID NO:94:	
30	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 38 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
35	(A) MAMERINA	
	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 20F/M13F primer</li></ul>	
40	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:	
	TGTAAAACGA CGGCCAGTCA CTGTGCCTGG CCTGATAC	38
45	(2) INFORMATION FOR SEQ ID NO:95:	
13	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 39 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	
50	<pre>(C) STRANDEDNESS: single (D) TOPOLOGY: linear</pre>	
55	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 20R/M13R primer</li><li>(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:</li></ul>	
	CAGGAAACAG CTATGACCAT GTTAAATTCA AAGTCTCTA	3.9
60	(2) INFORMATION FOR SEC ID NO:96:	

	5	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 39 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
	10	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 21F/M13F primer</li></ul>	
		(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:	
	15	TGTAAAACGA CGGCCAGTGG GTGTTTTATG CTTGGTTCT	39
		(2) INFORMATION FOR SEQ ID NO:97:	
-	20	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 40 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
:	25		
		<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 21R/M13R primer</li></ul>	
	30	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:	
:		CAGGAAACAG CTATGACCCA TTTCAACATA TTCCTTCCTG	40
	2.5	(2) INFORMATION FOR SEQ ID NO:98:	
	35	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 19 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	
	40	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
	4.5	(A) NAME/KEY: (B) LOCATION: (C) OFFICE AND ENDOMETION: 22F 1A primer	
	45	(D) OTHER INFORMATION: 22F-1A primer  (xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:	
		AACCACACC TTAAGATGA	19
	50	(2) INFORMATION FOR SEQ ID NO:99:	
		(i) SEQUENCE CHARACTERISTICS:	
	55	(A) LENGTH: 20 base pairs (B) TYPE: nucleic acid (C) STRANDEDNESS: single (D) TOPOLOGY: linear	
	60	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 22R-1A primer</li></ul>	

		(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:	
	5	GCATTAGTAG TGGATTTTGC	20
		(2) INFORMATION FOR SEQ ID NO:100:	
	10	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 16 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
	15		
		<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 23FII primer</li></ul>	
	20	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:	
		TCACTTCCAT TGCATC	16
	2.5	(2) INFORMATION FOR SEQ ID NO:101:	
the trees that he to the	25	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 17 base pairs</li><li>(B) TYPE: nucleic acid</li></ul>	
	30	<pre>(C) STRANDEDNESS: single (D) TOPOLOGY: linear</pre>	
	35	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 23RII primer</li></ul>	
		(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:	
	40	TGCCAACTGG TAGCTCC	17
	40	(2) INFORMATION FOR SEQ ID NO:102:	
	45	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
	50	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 24 2F primer</li></ul>	
		(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:	
	55	TACAGTTAGC AGCGACAAAA	20
		(2) INFORMATION FOR SEQ ID NO:103:	
	60	(i) SEQUENCE CHARACTERISTICS:  (A) LENGTH: 38 base pairs  (B) TYPE: pucleic acid	

	(C) STRANDEDNESS: single (D) TOPOLOGY: linear	
5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 24R/M13R primer</li></ul>	
10	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:	
	CAGGAAACAG CTATGACCAT TTGCCAACTG GTAGCTCC	38
1.5	(2) INFORMATION FOR SEQ ID NO:104:	
15	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li></ul>	
20	(D) TOPOLOGY: linear	
25	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 25F-7/23 primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:	
30	GCTTTCGCCA AATTCAGCTA	20
	(2) INFORMATION FOR SEQ ID NO:105:	
35	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
40	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 25R-7/23 primer</li></ul>	
4.5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:	
45	TACCAAAATG TGTGGTGATG	20
	(2) INFORMATION FOR SEQ ID NO:106:	
50	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 20 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
55		
60	(A) NAME/KEY: (B) LOCATION: (D) OTHER INFORMATION: 26-2F primer	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:	

(D) TOPOLOGY: linear

5	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 27BF/M13F primer</li></ul>	
	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:	
1.0	TGTAAAACGA CGGCCAGTGA ATTCTCCTCA GATGACTCCA	40
10	(2) INFORMATION FOR SEQ ID NO:111:	
15	<ul><li>(i) SEQUENCE CHARACTERISTICS:</li><li>(A) LENGTH: 38 base pairs</li><li>(B) TYPE: nucleic acid</li><li>(C) STRANDEDNESS: single</li><li>(D) TOPOLOGY: linear</li></ul>	
20	<ul><li>(A) NAME/KEY:</li><li>(B) LOCATION:</li><li>(D) OTHER INFORMATION: 27BR/M13R primer</li></ul>	
2.5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:	
25	CAGGAAACAG CTATGACCTC TTTGCTCATT GTGCAACA	38

20

25

# **WE CLAIM:**

5 1. A genomic DNA containing a BRCA2 gene,

wherein the first twelve nucleotides beginning exon 5 are 5'-

TCCTGTTGTTCT-3' as set forth in SEQ. ID. NO: 1,

wherein nucleotides numbers 5782-5790 are GTTTGTGTT as set forth in SEQ. ID. NO: 4, and

wherein the last 20 nucleotides ending exon 15 are 5'CTGCGTGTTCTCATAAACAG-3' as set forth in SEQ. ID. NO: 2 and the first 20
nucleotides beginning exon 16 are 5'-CTGTATACGTATGGCGTTTC-3' as set forth
in SEQ. ID. NO: 3.

2. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

1093 A 1342 A 1593 A 2457 T 2908 G 3199 A 3624 A 4035 T 7470 A 9079 G.

3. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

30 1093 A 1342 C 1593 A 2457 T 35 2908 G 3199 A 3624 A 4035 T 7470 A 9079 G.

4. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

	1093 A
	1342 C
5	1593 A
	2457 T
	2908 G
	3199 A
	3624 A
10	4035 C
	7470 A
	9079 G.

15 5. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

```
1342 A

20 1593 A

2457 C

2908 G

3199 G

3624 G

25 4035 T

7470 G

9079 G.
```

1093 C

6. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

```
1342 C
1593 A
1593 A
2457 T
2908 G
3199 A
3624 G
4035 T
40 7470 G
9079 G.
```

1093 A

7. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

```
1093 C
1342 C
1593 G
2457 C
50 2908 A
3199 G
```

20

3624 A
4035 T
7470 A
9079 A.

8. The genomic DNA according to claim 1 wherein the coding sequence nucleotides are as follows:

```
2024 C
4553 C
4815 G
5841 T
5972 C.
```

9. A DNA comprising a BRCA2 coding sequence, wherein nucleotide numbers 643-666 are

CTTAGTGAAAGTCCTGTTGTTCTA and

wherein nucleotides numbers 5782-5790 are GTTTGTGTT.

10. The DNA according to claim 9 wherein the coding sequence nucleotides are as follows:

```
25 1342 A
1593 A
2457 T
2908 G
3199 A
30 3624 A
4035 T
7470 A
9079 G.
```

1093 A

11. The DNA according to claim 9 wherein the coding sequence nucleotides are as follows:

```
1093 A
1342 C
40 1593 A
2457 T
2908 G
3199 A
3624 A
45 4035 T
7470 A
9079 G
as set forth in SEQ. ID. NO: 4.
```

12. The DNA according to claim 9 wherein the coding sequence nucleotides are as follows:

```
1093 A
1342 C
5 1593 A
2457 T
2908 G
3199 A
3624 A
10 4035 C
7470 A
9079 G
as set forth in SEQ. ID. NO: 6.
```

13. The DNA according to claim 9 wherein the coding sequence nucleotides are as follows:

```
1342 A
20 1593 A
2457 C
2908 G
3199 G
3624 G
25 4035 T
7470 G
9079 G
```

1093 C

as set forth in SEQ. ID. NO: 8.

14. The DNA according to claim 9 wherein the coding sequence nucleotides are as follows:

```
1342 C
1593 A
2457 T
2908 G
3199 A
3624 G
40 4035 T
7470 G
9079 G
as set forth in SEQ. ID. NO: 10.
```

1093 A

15. The DNA according to claim 9 wherein the coding sequence nucleotides are as follows:

```
1093 C
1342 C
50 1593 G
2457 C
```

25

```
2908 A
3199 G
3624 A
5 4035 T
7470 A
9079 A
as set forth in SEQ. ID. NO: 12.
```

10 16. The DNA according to claim 9 wherein the coding sequence nucleotides are as follows:

2024 C 4553 C 15 4815 G 5841 T 5972 C.

289

17. A BRCA2 protein having the following amino acids at the following peptide numbers:

372 histidine
894 valine
991 asparagine
1852 valine
1853 cysteine
1854 valine
2951 alanine

asparagine

as set forth in SEQ. ID. NO: 5.

18. The BRCA2 protein having the following amino acids at the following peptide numbers:

35 289 asparagine 372 asparagine 599 serine 894 valine 991 asparagine 40 2951 alanine.

19. The BRCA2 protein having the following amino acids at the following peptide numbers:

45 289 histidine
372 histidine
894 valine
991 asparatic acid
2951 alanine
50 as set forth in SEQ. ID. NO: 9.

- 20. The BRCA2 protein having the following amino acids at the following peptide numbers:
- 5 289 histidine
  - 372 asparagine
  - 894 isoleucine
  - 991 aspartic acid
  - 2951 threonine
- as set forth in SEQ. ID. NO: 13.
  - 21. The BRCA2 protein according to claims 17-20 having the following amino acids at the following peptide numbers:
- 15 **599** serine
  - 1442 serine
  - 1915 threonine.
  - 22. A haplotype of BRCA2 coding sequence (BRCA2° 1) as set forth in SEQ. ID.
- NO: 4 or a sequence complementary thereto.
  - 23. A BRCA2 protein comprising an amino acid sequence derived from BRCA2<sup>oml</sup> as set forth in SEQ. ID. NO: 5.
- 25 24. A haplotype of BRCA2 coding sequence (BRCA2<sup>om/2</sup>) as set forth in SEQ. ID. NO: 6 or a sequence complementary thereto.
  - 25. A BRCA2 protein comprising an amino acid sequence derived from BRCA2<sup>oml</sup>
    <sup>2</sup> as set forth in SEQ. ID. NO: 7.
  - 26. A haplotype of BRCA2 coding sequence (BRCA2° ) as set forth in SEQ. ID. NO: 8 or a sequence complementary thereto.
- 27. A BRCA2 protein comprising an amino acid sequence derived from BRCA2<sup>oml</sup> 35 3 as set forth in SEQ. ID. NO: 9.

15

20

25

- 28. A haplotype of BRCA2 coding sequence (BRCA2<sup>om 4</sup>) as set forth in SEQ. ID. NO: 10 or a sequence complementary thereto.
- <sup>4</sup> as set forth in SEQ. ID. NO: 11.
  - 30. A haplotype of BRCA2 coding sequence (BRCA2<sup>om, 5</sup>) as set forth in SEQ. ID. NO: 12 or a sequence complementary thereto.
  - 31. A BRCA2 protein comprising an amino acid sequence derived from BRCA2<sup>omi</sup> s set forth in SEQ. ID. NO: 13.
  - 32. A method of identifying individuals having a BRCA2 gene with a BRCA2 coding sequence not associated with disease, comprising:
    - (a) amplifying a DNA or a fragment thereof of an individual's BRCA2 coding sequence;
    - (b) sequencing said amplified DNA fragment;
    - (c) if necessary, repeating steps (a) and (b) until said individual's BRCA2 coding sequence is sufficiently sequenced to determine whether a mutation is present;
    - (d) comparing the sequence of said amplified DNA fragment to a
       BRCA2<sup>(omi)</sup> DNA sequence selecting from the group consisting of SEQ.
       ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID.
       NO: 12, and their respective complementary sequences;
    - (e) determining the presence of absence of each of the following polymorphic variations in said individual's BRCA2 coding sequence:
      - (i) AAT and CAT at position 1093,
      - (ii) CAT and AAT at position 1342,
      - (iii) TCA and TCG at position 1593,

10

15

25

- (iv) CAT and CAC at position 2457,
- (v) GTA and ATA at position 2908,
- (vi) AAC and GAC at position 3199,
- (vii) AAA and AAG at position 3624,
- (viii) GTT and GTC at position 4035,
- (ix) TCA and TCG at position 7470, and
- (x) GCC and ACC at position 9079; and
- (f) determining any sequence differences between said individual's BRCA2 coding sequences and a BRCA2<sup>(orm)</sup> DNA sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences, wherein the presence of said polymorphic variations and the absence of a variation outside of positions 1093, 1342, 1593, 2457, 2908, 3199, 3624, 4035, 7470, and 9079 is correlated with an absence of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA2 mutation in the BRCA2 coding sequence.
- 33. A method of identifying individuals having a BRCA2 gene with a BRCA2
   coding sequence not associated with disease, comprising:
  - (a) amplifying a DNA or a fragment thereof of an individual's BRCA2 coding sequence;
  - (b) sequencing said amplified DNA fragment;
  - (c) if necessary, repeating steps (a) and (b) until said individual's BRCA2 coding sequence is sufficiently sequenced to determine whether a mutation is present;
  - (d) comparing the sequence of said amplified DNA fragment to a BRCA2<sup>(orm)</sup> DNA sequence selecting from the group consisting of SEQ.
     ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences;

10

15

20

- (e) determining the presence of absence of each of the following polymorphic variations in said individual's BRCA2 coding sequence:
  - (i) AAT and CAT at position 1093,
  - (ii) CAT and AAT at position 1342,
  - (iii) TCA and TCG at position 1593,
  - (iv) CAT and CAC at position 2457,
  - (v) GTA and ATA at position 2908,
  - (vi) AAC and GAC at position 3199,
  - (vii) AAA and AAG at position 3624,
  - (viii) GTT and GTC at position 4035,
  - (ix) TCA and TCG at position 7470, and
  - (x) GCC and ACC at position 9079; and
- (f) determining any sequence differences between said individual's BRCA2 coding sequences and a BRCA2<sup>(orm)</sup> DNA sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences, wherein the presence of said polymorphic variations and the absence of a variation outside of positions 1093, 1342, 1593, 2457, 2908, 3199, 3624, 4035, 7470, and 9079 is correlated with an absence of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA2 mutation in the BRCA2 coding sequence; wherein, codon variations occur at the following frequencies, respectively, in a Caucasian population of individuals free of disease:
  - (i) at position 1093, AAT and CAT occur at frequencies from about 75-85%, and from about 15-25%, respectively,
  - (ii) at position 1342, CAT and AAT occur at frequencies from about 35-45%, and from about 55-65%, respectively.
  - (iii) at position 1593, TCA and TCG occur at frequencies from about 85-95%, and from about 5-15%, respectively,

25

\_ \_

25

(iv) at position 2457, CAT and CAC occur at frequencies from about 75-85%, and from about 15-25%, respectively, at position 2908, GTA and ATA occur at frequencies from (v) 5 about 85-95%, and from about 5-15%, respectively. (vi) at position 3199, AAC and GAC occur at frequencies from about 75-85%, and from about 15-25%. respectively. (vii) at position 3624, AAA and AAG occur at frequencies 10 from about 75-85%, and from about 15-25%, respectively, (viii) at position 4035, GTT and GTC occur at frequencies from about 85-95%, and from about 5-15%, respectively, (ix) at position 7470, TCA and TCG occur at frequencies from

- 34. A method of detecting an increased genetic susceptibility to breast and ovarian cancer in an individual resulting from the presence of a mutation in the BRCA2 coding sequence, comprising:
  - (a) amplifying a DNA or a fragment thereof of an individual's BRCA2 coding sequence;

about 75-85%, and from about 15-25%, respectively, and

from about 85-95%, and from about 5-15%, respectively.

at position 9079, GCC and ACC occur at frequencies

(b) sequencing said amplified DNA fragment;

(x)

- (c) if necessary, repeating steps (a) and (b) until said individual's BRCA2 coding sequence is sufficiently sequenced to determine whether a mutation is present;
- (d) comparing the sequence of said amplified DNA fragment to a BRCA2<sup>(omi)</sup> DNA sequence selected from the group consisting of SEQ.

10

15

20

- ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences;
- (e) determining any sequence differences between said individual's BRCA2 coding sequences and a BRCA2<sup>(orm)</sup> DNA sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences in order to determine the presence or absence of base changes in said individual's BRCA2 coding sequence wherein a base change which is not any one of the following:
  - (i) AAT and CAT at position 1093,
  - (ii) CAT and AAT at position 1342,
  - (iii) TCA and TCG at position 1593,
  - (iv) CAT and CAC at position 2457,
  - (v) GTA and ATA at position 2908,
  - (vi) AAC and GAC at position 3199,
  - (vii) AAA and AAG at position 3624.
  - (viii) GTT and GTC at position 4035.
  - (ix) TCA and TCG at position 7470, and
  - (x) GCC and ACC at position 9079, is correlated with the potential of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA2 mutation in the BRCA2 coding sequence.
- 35. A method of detecting an increased genetic susceptibility to breast and ovarian cancer in an individual resulting from the presence of a mutation in the BRCA2 coding sequence, comprising:
  - (a) amplifying a DNA or a fragment thereof of an individual's BRCA2 coding sequence;
  - (b) sequencing said amplified DNA fragment;

- (c) if necessary, repeating steps (a) and (b) until said individual's BRCA2 coding sequence is sufficiently sequenced to determine whether a mutation is present;
- (d) comparing the sequence of said amplified DNA fragment to a BRCA2<sup>(orni)</sup> DNA sequence selected from the group consisting of: SEQ.
   ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences;
- (e) determining any sequence differences between said individual's BRCA2 coding sequences and a BRCA2<sup>(orni)</sup> DNA sequence selected from the group consisting of: SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, and their respective complementary sequences in order to determine the presence or absence of base changes in said individual's BRCA2 coding sequence wherein a base change which is not any one of the following:
  - (i) AAT and CAT at position 1093,
  - (ii) CAT and AAT at position 1342,
  - (iii) TCA and TCG at position 1593,
  - (iv) CAT and CAC at position 2457,
  - (v) GTA and ATA at position 2908,
  - (vi) AAC and GAC at position 3199,
  - (vii) AAA and AAG at position 3624,
  - (viii) GTT and GTC at position 4035,
  - (ix) TCA and TCG at position 7470, and
  - (x) GCC and ACC at position 9079, is correlated with the potential of increased genetic susceptibility to breast or ovarian cancer resulting from a BRCA2 mutation in the BRCA2 coding sequence, wherein, codon variations occur at the following frequencies, respectively, in a Caucasian population of individuals free of disease:

15

20

- (i) at position 1093, AAT and CAT occur at frequencies from about 75-85%, and from about 15-25%, respectively. (ii) at position 1342, CAT and AAT occur at frequencies from 5 about 35-45%, and from about 55-65%, respectively, (iii) at position 1593, TCA and TCG occur at frequencies from about 85-95%, and from about 5-15%, respectively, (iv) at position 2457, CAT and CAC occur at frequencies from about 75-85%, and from about 15-25%, respectively. 10 (v) at position 2908, GTA and ATA occur at frequencies from about 85-95%, and from about 5-15%, respectively, (vi) at position 3199, AAC and GAC occur at frequencies from about 75-85%, and from about 15-25%, respectively, at position 3624, AAA and AAG occur at frequencies 15 (vii) from about 75-85%, and from about 15-25%, respectively, at position 4035, GTT and GTC occur at frequencies from (viii) about 85-95%, and from about 5-15%, respectively, at position 7470, TCA and TCG occur at frequencies from 20 (ix) about 75-85%, and from about 15-25%, respectively, and at position 9079, GCC and ACC occur at frequencies (x)
- 36. A method according to any of the claims 32-35 wherein the said amplifying is performed by annealing at least one oligonucleotide primer to said DNA fragment and extending the oligonucleotide primer by an agent for polymerization.

from about 85-95%, and from about 5-15%, respectively.

37. A method according to claim 36 wherein said oligonucleotide primer is directly or indirectly labeled with a radioactive label, a fluorescent label, a bioluminescent label, a chemiluminescent label, a metal chelator, or an enzyme label.

38. A BRCA2 coding sequence according to claims 32, wherein the codon pairs occur at the following frequencies:

- (i) at position 1093, AAT and CAT occur at frequencies from about 75-85%, and from about 15-25%, respectively,
- (ii) at position 1342, CAT and AAT occur at frequencies from about 35-45%, and from about 55-65%, respectively,
- (iii) at position 1593, TCA and TCG occur at frequencies from about 85-95%, and from about 5-15%, respectively,
- (iv) at position 2457, CAT and CAC occur at frequencies from about 75-85%, and from about 15-25%, respectively,
- (v) at position 2908, GTA and ATA occur at frequencies from about 85-95%, and from about 5-15%, respectively,
- (vi) at position 3199, AAC and GAC occur at frequencies from about 75-85%, and from about 15-25%, respectively,
- (vii) at position 3624, AAA and AAG occur at frequencies from about 75-85%, and from about 15-25%, respectively,
- (viii) at position 4035, GTT and GTC occur at frequencies from about 85-95%, and from about 5-15%, respectively,
- (ix) at position 7470, TCA and TCG occur at frequencies from about 75-85%, and from about 15-25%, respectively, and
- (x) at position 9079, GCC and ACC occur at frequencies from about 85-95%, and from about 5-15%, respectively.

10

5

15

20

10

15

20

30

- 39. An oligonucleotide primer capable of hybridizing to a sample of BRCA2 gene, or its respective complementary sequences selected from the group consisting of SEQ. ID. NO: 14, 19, 22, 23, 25, 26, 29-76, 83, 85-88, 90, 91, 97, 98, 101, and 104-107.
- 40. A chip array having "n" elements for performing allele specific sequence-based techniques comprising a solid phase chip and oligonucleotides having "n" different nucleotide sequences,

wherein "n" is an interger greater than or equal to ten,

wherein said oligonucleotides are bound to said solid phase chip in a manner which permits said oligonucleotides to effectively hybridize to complementary oligonucleotides or polynucleotides,

wherein oligonucleotides having different nucleotide sequence are bound to said solid phase chip at different locations so that a particular location on said solid phase chip exclusively binds oligonucleotides having a specific nucleotide sequence, and

wherein at least ten oligonucleotides are capable of specifically hybridizing to the BRCA2 DNA having the sequence as set forth in SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 10, SEQ. ID. NO: 12 or their respective complementary sequences, at least one oligonucleotide being capable of specifically hybridizing at each of the nucleotide positions 1093, 1342, 1593, 2457, 2908, 3199, 3624, 4035, 7470, 9079, or complementary thereto.

- 25 41. A method of performing gene therapy on a patient, comprising:
  - a) contacting cancer cells *in vivo* with an effective amount of a vector comprising DNA containing at least a portion of BRCA2 sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, or their respective complementary sequences
    - b) allowing the vector to enter the cancer cells, and
    - c) measuring a reduction in tumor growth.
  - 42. The method according to claim 41 wherein said cancer cells have a mutation in the BRCA2 gene.

25

5

10

- 43. The method according to claim 41 wherein said patient has a mutation in the BRCA2 gene of non-cancer cells.
- 44. A method of performing gene therapy on a patient or a sample, comprising:
- a) contacting cells *in vivo* or *in vitro* with an effective amount of a vector comprising DNA containing at least a portion of BRCA2 sequence selected from the group consisting of SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO:
- 10, SEQ. ID. NO: 12, or their respective complementary sequences, and
  - b) allowing the vector to enter the cells,

wherein said patient has a reduced susceptibility for developing a cancer associated with a mutation in the BRCA2 gene.

- 15 45. A method according to claim 44 wherein said cells include healthy breast, ovarian or pancreatic tissues.
  - 46. A method according to claim 44 wherein a patient has an inherited mutation in the BRCA2 gene.
  - 47. A method of treating a patient suspected of having a tumor, comprising:
  - a) administering to a patient an effective amount of BRCA2 tumor growth inhibitor having an amino acid sequence selected from the group consisting of SEQ. ID. NO: 5, SEQ. ID. NO: 7, SEQ. ID. NO: 9, SEQ. ID. NO: 11, SEQ. ID. NO: 13, any fragments thereto, and any functional equivalent thereof;
    - b) allowing the patient's cells to take up the protein, and
    - c) measuring a reduction in tumor growth.
- 48. The method according to claim 47 wherein said tumor is a breast cancer, an ovarian cancer or a pancreatic cancer.
  - 49. The method according to claim 47 wherein said patient has an inherited mutation in the BRCA2 gene.

15

20

- 50. A method of preventing the formation or growth of a tumor, comprising:
- a) adminstering to a patient an effective amount of BRCA2 tumor growth inhibiting protein having an amino acid sequence selected from the group consisting of SEQ. ID. NO: 5, SEQ. ID. NO: 7, SEQ. ID. NO: 9, SEQ. ID. NO: 11, SEQ. ID. NO: 13, any fragments thereto, and any functional equivalent thereof; and
  - b) allowing the patient cells to take up the protein.
- 51. The method according to claim 31 wherein the protein is administered parenternally, by buccal adsorption or inhalation.
  - 52. A cloning vector comprising:
  - (a) a DNA sequence as set forth in SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, or any fragments thereof; and
  - (b) one or more suitable regulatory sequences to induce replication and/or integration in a host cell.
  - 53. An expression vector comprising a DNA sequence as set forth in SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO: 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, or any fragments thereof operatively linked to one or more promoter sequences capable of directing expression of said sequence in a host cell.
  - 54. A host cell transformed with the vector according to claim 52 or 53.
- 55. A BRCA2 polypeptide which is selected from the group consisting of:

  (a) a fragment of BRCA2 protein sequence as set forth in SEQ. ID. NO: 5,
  SEQ. ID. NO: 7, SEQ. ID. NO: 9, SEQ. ID. NO: 11, or SEQ. ID. NO:13;
  (b) an amino acid sequence which is substantially homologous to the BRCA2 protein sequence as set forth in SEQ. ID. NO: 5, SEQ. ID. NO: 7, SEQ. ID.

  30 NO: 9, SEQ. ID. NO: 11, or SEQ. ID. NO: 13;
  (c) a molecule which has similar function to the BRCA2 protein; and

(d) a fusion protein of (a), (b), or (c).

- 56. An anti-BRCA2 antibody wherein a molecule according to claims 17-21, 23, 25, 27, 29, 31, or 55 is used as an immunogen.
- 5 57. A diagnostic reagent comprising a molecule selected from the group consisting of:
  - (a) a DNA sequence as set forth in SEQ. ID. NO: 4, SEQ. ID. NO: 6, SEQ. ID. NO:
  - 8, SEQ. ID. NO: 10, SEQ. ID. NO: 12, or their complementary sequences;
  - (b) a nucleic acid fragment of (a) comprising at least 10 nucleotide in length;
- 10 (c) a sequence which hybridizes to (a) or (b);
  - (d) a polypeptide according to claim 17-21, 23, 25, 27, 29, 31, or 55; and
  - (e) an antibody which specifically binds to the polypeptide of (d).
  - 58. A pharmaceutical composition comprising a molecule according to any one of the claims 17-21, 23, 25, 27, 29, 31, 55 in a pharmaceutically acceptable carrier.
    - 59. A pharmaceutical composition comprising a molecule according claim 56 in a pharmaceutically acceptable carrier.
- 20 60. A pharmaceutical composition comprising a molecule according to claim 57 in a pharmaceutically acceptable carrier.

10

# ABSTRACT OF THE DISCLOSURE

Five novel DNA and protein sequences have been determined for the BRCA2 gene, as have been ten polymorphic sites and their rates of occurrence in the normal alleles of BRCA2. The sequences BRCA2<sup>(orm; 1-5)</sup> and the ten polymorphic sites will provide greater accuracy and reliability for genetic testing. One skilled in the art will be better able to avoid misinterpretations of changes in the gene and/or protein sequence, determine the presence of a normal sequence, and of mutations of BRCA2. This invention is also related to a method of performing gene therapy with BRCA2<sup>(orm; 1-5)</sup> coding sequences or fragments thereof. This invention is further related to protein therapy with BRCA2<sup>(orm; 1-5)</sup> proteins or their functional equivalents.

# Figure 1A

Exon 2

taagtgcattttggtcttctgttttgcagACTTATTTACCAAGCATTGGAGGAATATCGTAGGTAAAA ATĞCCTATTGGATCCAAAGAGAGGCCAACATTTTTTGAAATTTTTAAGACACGCTGC AACAAAGCAGgtattgacaaattttatataac

Exon 3

gggattttttttttaaatagATTTAGGACCAATAAGTCTTAATTGGTTTGAAGAACTTTCTTCAG ĂĂĞCTCCACCCTATAATTCTGAACCTGCAGAAGAATCTGAACATAAAAACAACAATT ACGAACCAAACCTATTTAAAACTCCACAAAGGAAACCATCTTATAATCAGCTGGCTT CAACTCCAATAATATTCAAAGAGCAAGGGCTGACTCTGCCGCTGTACCAATCTCCT GTAAAAGAATTAGATAAATTCAAATTAGACTTAGgtaagtaatgcaatatggtagactgggg

Exon 4

 $t cact ga att att gtact gttt cag {\tt GAAGGAATGTTCCCAATAGTAGACATAAAAGTCTTCGCACA}$ GTGAAAACTAAAATGGATCAAGCAGATGATGTTTCCTGTCCACTTCTAAATTCTTGT CTTAGTGAAAGgtatgatgaagctattatattaaaa

Exon 5

agggatttgctttgtTTTATTTTAGTCCTGTTGTTCTACAATGTACACATGTAACACCACAAA GĂĞATĂAGTCAGgtatgattaaaaacaatgctttttattctt

Exon 6

ttaacaattttcccctttttttacccccagTGGTATGTGGGAGTTTGTTTCATACACCAAAGTTTGTG **AAGgtaaatatt** 

Exon 7

TCTGAAĂĞTCTAGGAGCTGAGGTGGATCCTGĂTATGTCTTGGTCAAGTTCTTTAGC TACACCACCCACCCTTAGTTCTACTGTGCTCATAGgtaataata

Exon 8

ttttatcttacagTCAGAAATGAAGAAGCATCTGAAACTGTATTTCCTCATGATACTACTGC Tgtaagtaaatatgacattgattagact

Exon 9

taaactataatttttgcagAATGTGAAAAGCTATTTTTCCAATCATGATGAAAGTCTGAAGAAA AATGATAGAŤTŤATCGCTTCTGTGACAGACAGTGAAAACACAAATCAAAGAGAAGC TGCAAGTCATGgtaagtcctct

Exon 10

tta at gtg ctt ctg tttt at a cttta a cag GATTTGGAAAAACATCAGGGAATTCATTTAAAGTAAATAGCTĞCAAĂGACCACATTĞGAAAGTCAATGCCAAATGTCCTAGAAGATĢAAGTATAT GAAACAGTTGTAGATACCTCTGAAGAAGATAGTTTTTCATTATGTTTTTCTAAATGTA GAACAAAAATCTACAAAAAGTAAGAACTAGCAAGACTAGGAAAAAAATTTTCCATG TTGTATCTGAAGTGGAACCAAATGATACTGATCCATTAGATTCAAATGTAGCAAATC

# Figure 1B

# Exon 11

tttgtgtttttatgtttagGTTTATTGCATTCTTCTGTGAAAAGAAGCTGTTCACAGAATGATTCT GĂĂGAAČCAĂCTTTGTCCTTAACTAGCTCTTTTGGGACAATTCTGAGGAAATGTTCT AGAAATGAAACATGTTCTAATAATACAGTAATCTCTCAGGATCTTGATTATAAAGAA GCAAAATGTAATAAGGAAAAACTACAGTTATTTATTACCCCAGAAGCTGATTCTCTG TCATGCCTGCAGGAAGGACAGTGTGAAAATGATCCAAAAAGCAAAAAGTTTCAGA TATAAAAGAAGAGGTCTTGGCTGCAGCATGTCACCCAGTACAACATTCAAAAGTGG AATACAGTGATACTGACTTTCAATCCCAGAAAAGTCTTTTATATGATCATGAAAATG CCAGCACTCTTATTTTAACTCCTACTTCCAAGGATGTTCTGTCAAACCTAGTCATGA TTTCTAGAGGCAAAGAATCATACAAAATGTCAGACAAGCTCAAAGGTAACAATTATG CTTTAAATGAAAATTATAAAAACGTTGAGCTGTTGCCACCTGAAAAATACATGAGAG TAGCATCACCTTCAAGAAAGGTACAATTCAACCAAAACACAAATCTAAGAGTAATCC AAAAAAATCAAGAAGAAACTACTTCAATTTCAAAAATAACTGTCAATCCAGACTCTG AAGAACTTTCTCAGACAATGAGAATAATTTTGTCTTCCAAGTAGCTAATGAAAGGA ATAATCTTGCTTTAGGAAATACTAAGGAACTTCATGAAACAGACTTGACTTGTGTAA ACGAACCCATTTTCAAGAACTCTACCATGGTTTTATATGGAGACACAGGTGATAAAC AAGCAACCCAAGTGTCAATTAAAAAAGATTTGGTTTATGTTCTTGCAGAGGAGAAC AAAAATAGTGTAAAGCAGCATATAAAAATGACTCTAGGTCAAGATTTAAAATCGGAC ATCTCCTTGAATATAGATAAAATACCAGAAAAAAAATAATGATTACATGAACAAATGG GCAGGACTCTTAGGTCCAATTTCAAATCACAGTTTTGGAGGTAGCTTCAGAACAGC TTCAAATAAGGAAATCAAGCTCTCTGAACATAACATTAAGAAGAGCAAAATGTTCTT CAAAGATATTGAAGAACAATATCCTACTAGTTTAGCTTGTGTGAAATTGTAAATAC CTTGGCATTAGATAATCAAAAGAAACTGAGCAAGCCTCAGTCAATTAATACTGTATC TGCACATTTACAGAGTAGTGTAGTTGTTCTGATTGTAAAAATAGTCATATAACCCC TCAGATGTTATTTCCAAGCAGGATTTTAATTCAAACCATAATTTAACACCTAGCCAA AAGGCAGAAATTACAGAACTTTCTACTATATTAGAAGAATCAGGAAGTCAGTTTGAA TTTACTCAGTTTAGAAAACCAAGCTACATATTGCAGAAGAGTACATTTGAAGTGCCT GAAAACCAGATGACTATCTTAAAGACCACTTCTGAGGAATGCAGAGATGCTGATCT 

# Figure 1C

AAGGTACAGTTGAAATTAAACGGAAGTTTGCTGGCCTGTTGAAAAATGACTGTAAC AAAAGTGCTTCTGGTTATTTAACAGATGAAAATGAAGTGGGGTTTAGGGGGCTTTTAT TCTGCTCATGGCACAAAACTGAATGTTTCTACTGAAGCTCTGCAAAAAGCTGTGAA ACTGTTTAGTGATATTGAGAATATTAGTGAGGAAACTTCTGCAGAGGTACATCCAAT AAGTTTATCTTCAAGTAAATGTCATGATTCTGTTGTTTCAATGTTTAAGATAGAAAAT CATAATGATAAACTGTAAGTGAAAAAAATAATAAATGCCAACTGATATTACAAAATA ATATTGAAATGACTACTGGCACTTTTGTTGAAGAAATTACTGAAAATTACAAGAGAA ATACTGAAAATGAAGATAACAAATATACTGCTGCCAGTAGAAATTCTCATAACTTAG AATTTGATGGCAGTGATTCAAGTAAAAATGATACTGTTTGTATTCATAAAGATGAAA CGGACTTGCTATTTACTGATCAGCACAACATATGTCTTAAATTATCTGGCCAGTTTA TGAAGGAGGAAACACTCAGATTAAAGAAGATTTGTCAGATTTAACTTTTTTGGAAG TTGCGAAAGCTCAAGAAGCATGTCATGGTAATACTTCAAATAAAGAACAGTTAACT GCTACTAAAACGGAGCAAAATATAAAAGATTTTGAGACTTCTGATACATTTTTTCAG ACTGCAAGTGGGAAAAATATTAGTGTCGCCAAAGAGTCATTTAATAAAATTGTAAAT TTCTTTGATCAGAAACCAGAAGAATTGCATAACTTTTCCTTAAATTCTGAATTACATT CTGACATAAGAAAGAACAAAATGGACATTCTAAGTTATGAGGAAACAGACATAGTT AAACACAAAATACTGAAAGAAGTGTCCCAGTTGGTACTGGAAATCAACTAGTGAC CTTCCAGGGACACCCGAACGTGATGAAAAGATCAAAGAACCTACTCTGTTGGGTT TTCATACAGCTAGCGGGAAAAAAGTTAAAATTGCAAAGGAATCTTTGGACAAAGTG AAAAACCTTTTTGATGAAAAAGAGCAAGGTACTAGTGAAATCACCAGTTTTAGCCAT CAATGGGCAAAGACCCTAAAGTACAGAGAGGCCTGTAAAGACCTTGAATTAGCAT GTGAGACCATTGAGATCACAGCTGCCCCAAAGTGTAAAGAAATGCAGAATTCTCTC AATAATGATAAAAACCTTGTTTCTATTGAGACTGTGGTGCCACCTAAGCTCTTAAGT GATAATTTATGTAGACAAACTGAAAATCTCAAAAACATCAAAAAGTATCTTTTTGAAAG TTAAAGTACATGAAAATGTAGAAAAAGAAACAGCAAAAAGTCCTGCAACTTGTTACA CAAATCAGTCCCCTTATTCAGTCATTGAAAATTCAGCCTTAGCTTTTTACACAAGTT GTAGTAGAAAACTTCTGTGAGTCAGACTTCATTACTTGAAGCAAAAAAATGGCTTA GAGAAGGAATATTTGATGGTCAACCAGAAAGAATAAATACTGCAGATTATGTAGGA AATTATTTGTATGAAAATAATTCAAACAGTACTATAGCTGAAAATGACAAAAATCATC TCTCCGAAAAACAAGATACTTATTTAAGTAACAGTAGCATGTCTAACAGCTATTCCT ACCATTCTGATGAGGTATATAATGATTCAGGATATCTCTCAAAAAATAAACTTGATT AAGTAATATCCAATGTAAAAGATGCAAATGCATACCCACAAACTGTAAATGAAGATA TTTGCGTTGAGGAACTTGTGACTAGCTCTTCACCCTGCAAAAATAAAAATGCAGCC ATTAAATTGTCCATATCTAATAGTAATAATTTTGAGGTAGGGCCACCTGCATTTAGG ATAGCCAGTGGTAAAATCGTTTGTGTTTCACATGAAACAATTAAAAAAGTGAAAGAC ATATTTACAGACAGTTTCAGTAAAGTAATTAAGGAAAACAACGAGAATAAATCAAAA ATTTGCCAAACGAAAATTATGGCAGGTTGTTACGAGGCATTGGATGATTCAGAGGA TATTCTTCATAACTCTCTAGATAATGATGAATGTAGCACGCATTCACATAAGGTTTTT GCTGACATTCAGAGTGAAGAAATTTTACAACATAACCAAAATATGTCTGGATTGGA GAAAGTTTCTAAAATATCACCTTGTGATGTTAGTTTGGAAACTTCAGATATATGTAAA TGTAGTATAGGGAAGCTTCATAAGTCAGTCTCATCTGCAAATACTTGTGGGATTTTT AGCACAGCAAGTGGAAAATCTGTCCAGGTATCAGATGCTTCATTACAAAACGCAAG ACAAGTGTTTTCTGAAATAGAAGATAGTACCAAGCAAGTCTTTTCCAAAGTATTGTT TAAAAGTAACGAACATTCAGACCAGCTCACAAGAGAAAAATACTGCTATACGTA CTCCAGAACATTTAATATCCCAAAAAGGCTTTTCATATAATGTGGTAAATTCATCTG

# Figure 1D

CTTTCTCTGGATTTAGTACAGCAAGTGGAAAGCAAGTTTCCATTTTAGAAAGTTCCT TACACAAAGTTAAGGGAGTGTTAGAGGAATTTGATTTAATCAGAACTGAGCATAGT CTTCACTATTCACCTACGTCTAGACAAAATGTATCAAAAATACTTCCTCGTGTTGAT AAGAGAAACCCAGAGCACTGTGTAAACTCAGAAATGGAAAAAACCTGCAGTAAAGA ATTTAAATTATCAAATAACTTAAATGTTGAAGGTGGTTCTTCAGAAAATAATCACTCT ATTAAAGTTTCTCCATATCTCTCAATTTCAACAAGACAACAACAGTTGGTATTAG GAACCAAAGTCTCACTTGTTGAGAACATTCATGTTTTGGGAAAAGAACAGGCTTCA CCTAAAAACGTAAAAATGGAAATTGGTAAAACTGAAACTTTTTCTGATGTTCCTGTG AAAACAAATATAGAAGTTTGTTCTACTTACTCCAAAGATTCAGAAAACTACTTTGAAA CAGAAGCAGTAGAAATTGCTAAAGCTTTTATGGAAGATGATGAACTGACAGATTCT AAACTGCCAAGTCATGCCACACATTCTCTTTTTACATGTCCCGAAAATGAGGAAATG GTTTTGTCAAATTCAAGAATTGGAAAAAGAAGAGGGGGGGCCCCTTATCTTAGTGGgt aagtgttcatttttacctttcgtgttgccaatca

# Exon 12

aaaacatatatgaaatatttctttttagGAGAACCCTCAATCAAAAGAAACTTATTAAATGAATTTG ACAGGATAATAGAAAATČAAGAAAAATCCTTAAAGGCTTCAAAAAGCACTCCAGAT Ggtaaaattagctttttatttata

# Exon 13

aatatgtaatataaaataattgtttcctagGCACAATAAAAGATCGAAGATTGTTTATGCATCATGT TTCTTTAGAGCCGATTACCTGTGTACCCTTTCGgtaagacatgtttaaatttttctaa

## Exon 14

cccattgcagCACAACTAAGGAACGTCAAGAGATACAGAATCCAAATTTTACCGCACC TGGTČAAĞAATTTCTGTCTAAATCTCATTTGTATGAACATCTGACTTTGGAAAAATCT TCAAGCAATTTAGCAGTTTCAGGACATCCATTTTATCAAGTTTCTGCTACAAGAAAT GAAAAAATGAGACACTTGATTACTACAGGCAGACCAAACCAAAGTCTTTGTTCCACC TTTTAAAACTAAATCaCATTTTCACAGAGTTGAACAGTGTGTTAGGAATATTAACTTG GAGGAAACAGACAAAGCAAAACATTGATGGACATGGCTCTGATGATAAAAA TAAGATTAATGACAATGAGATTCATCAGTTTAACAAAAACAACTCCAATCAAGCAGC AGCTGTAACTTTCACAAAGTGTGAAGAAGAACCTTTAGgtattgtatgacaatttgtgtgatgaatt

#### Exon 15

tttttgctaagtatttattctttgatagATTTAATTACAAGTCTTCAGAATGCCAGAGATATACAGGAT ATGCGAATTAAGAAGAAACAAAGGCAACGCGTCTTTCCACAGCCAGGCAGTCTGTA TCTTGCAAAAACATCCACTCTGCCTCGAATCTCTCTGAAAGCAGCAGTAGGAGGCC  ${\sf AAGTTCCCTCTGCgtgtccccataaacaggtatgtgt}$ 

#### Exon 16

tttttcttttttgtgtgtgtttattttgtgtag GTGTTCTCATAAACAG CTGTATACGTATGGCGTTTCTAAACATTGCATAAAAATTAACAGCAAAAATGCAGAGTCTTTTCAGTTTCACACTGAAGA TTATTTTGGTAAGGAAAGTTTATGGACTGGAAAAGGAATACAGTTGGCTGATGGTG GATGGCTCATACCCTCCAATGATGGAAAGGCTGGAAAAGAAGAATTTTATAGgtactct atgcaaaaagattgtgtgttaacttttatg

# Figure 1E

# Exon 17

ttatttgttcagGGCTCTGTGACACTCCAGGTGTGGATCCAAAGCTTATTTCTAGAATTTGGGTTTATAATCACTATAGATGGATCATATGGAAACTGGCAGCTATGGAATGTGCCTTTCCTAAGGAATTTGCTAATAGATGCCTAAGCCCAGAAAGGGTGCTTCTTCAACTAAATACAGgcaagtttaaagcatt

#### Exon 18

ttttgttttcacttttagaTATGATACGGAAATTGATAGAAGCAGAAGATCGGCTATAAAAAAGA TAATGGAAAGGGATGACACAGCTGCAAAAACACTTGTTCTCTGTGTTTCTGACATA ATTTCATTGAGCGCAAATATATCTGAAACTTCTAGCAATAAAACTAGTAGTGCAGAT ACCCAAAAAGTGGCCATTATTGAACTTACAGATGGGTGGTATGCTGTTAAGGCCCA GTTAGATCCTCCCCTCTTAGCTGTCTTAAAGAATGGCAGACTGACAGTTGGTCAGA AGATTATTCTTCATGGAGCAGAACTGGTGGGCTCTCCTGATGCCTGTACACCTCTT GAAGCCCCAGAATCTCTTATGTTAAAGgtaaatt

## Exon 19

taaatcaatatatttattaatttgtccagATTTCTGCTAACAGTACTCGGCCTGCTCGCTGGTATAC CAAACTTGGATTCTTTCCTGACCCTAGACCTTTTCCTCTGCCCTTATCATCGCTTTT CAGTGATGGAGGAAATGTTGGTTGTTGATGTAATTATTCAAAGAGCATACCCTAT ACAGgtatgatgtattcttgaaactta

#### Exon 20

tttggtgtgtgtaacacattattacagTGGATGGAGAAGACATCATCTGGATTATACATATTTCGC AATGAAAGAGGAGGAAGAAAGGAAGCAGCAAAATATGTGGAGGCCCAACAAAAGA GACTAGAAGCCTTATTCACTAAAATTCAGGAGGAATTTGAAGAACATGAAGgtaaaatt agttatatggtacacattgttatttc

#### Exon 21

agtitagtgaattaataatccttttgtttcttagAAAACACAACAAAACCATATTTACCATCACGTGCAC TAACAAGACAGCAGCTCGTGCTTTGCAAGATGGTGCAGAGCTTTATGAAGCAGTG AAGAATGCAGCAGACCCAGCTTACCTTGAGgtgagagagtaagaggacatataatgag

#### Exon 22

## Exon 23

tctccaaacagTTATACTGAGTATTTGGCGTCCATCATCAGATTTATATTCTCTGTTAACA GAAGGAAAGAGATACAGAATTTATCATCTTGCAACTTCAAAATCTAAAAGTAAATCT GAAAGAGCTAACATACAGTTAGCAGCGACAAAAAAACTCAGTATCAACAACTACC Ggtacaaacctttcattgtaattttt

# Figure 1F

#### Exon 24

## Exon 25

taacattcttttctttttttttccattctagGACTTGCCCCTTTCGTCTATTTGTCAGACGAATGTTACAA TTTACTGGCAATAAAGTTTTGGATAGACCTTAATGAGGACATTATTAAGCCTCATAT GTTAATTGCTGCAAGCAACCTCCAGTGGCGACCAGAATCCAAATCAGGCCTTCTTA CTTTATTTGCTGGAGATTTTTCTGTGTTTTCTGCTAGTCCAAAAGAGGGCCACTTTC AAGAGACATTCAACAAAATGAAAAATACTGTTGAGgtaaggtta

## Exon 26

ataaagcagcttttccacttatttcttagAATATTGACATACTTTGCAATGAAGCAGAAAACAAGCT TATGCATATACTGCATGCAAATGATCCCAAGTGGTCCACCCCAACTAAAGACTGTA CTTCAGGGCCGTACACTGCTCAAATCATTCCTGGTACAGGAAACAAGCTTCTGgtaa gttaatgtaaactcaaggaatattataag

#### Exon 27

# Figure 2A

#### Exon 2

taagtgcattttggtcttctgttttgcagACTTATTTACCAAGCATTGGAGGAATATCGTAGGTAAAA <u>ATG</u>CCTATTGGATCCAAAGAGGGCCAACATTTTTTGAAATTTTTAAGACACGCTGC AACAAAGCAGgtattgacaaattttatataac

#### Fxon 3

gggatttttttttaaatagATTTAGGACCAATAAGTCTTAATTGGTTTGAAGAACTTTCTTCAGAAGCTCCACCCTATAATTCTGAACCTGCAGAAGAATCTGAACATAAAAACAACAATTACGAACCAAACCTATTTAAAACTCCACAAAGGAAACCATCTTATAATCAGCTGGCTTCAACTCCAAATAATATTCAAAGAGCAAGGGCTGACTCTGCCGCTGTACCAATCTCCTGTAAAAAGAATTAGATAAATTCAAATTAGACTTAGGtaagtaatgcaatatggtagactgggg

#### Exon 4

tcactgaattattgtactgtttcagGAAGGAATGTTCCCAATAGTAGACATAAAAGTCTTCGCACA GTGAAAACTAAAATGGATCAAGCAGATGATGTTTCCTGTCCACTTCTAAATTCTTGT CTTAGTGAAAGgtatgatgaagctattatattaaaa

#### Exon 5

agggatttgctttgttttattttagTCCTGTTGTTCTACAATGTACACATGTAACACCACAAAGAGATAAAGTCAGgtatgattaaaaacaatgctttttattctt

#### Exon 6

ttaacaattttcccctttttttacccccagTGGTATGTGGGAGTTTGTTTCATACACCAAAGTTTGTGAAAGtaaatatt

#### Exon 7

## Exon 8

ttttatcttacagTCAGAAATGAAGAAGCATCTGAAACTGTATTTCCTCATGATACTACTGC Tgtaagtaaatatgacattgattagact

#### Exon 9

taaactataatttttgcagAATGTGAAAAGCTATTTTTCCAATCATGATGAAAGTCTGAAGAAA AATGATAGATTTATCGCTTCTGTGACAGACAGTGAAAACACAAATCAAAGAGAAGC TGCAAGTCATGgtaagtcctct

## Exon 10

# Figure 2B

AGAAGCCCTTTGAGAGTGGAAGTGACAAAATCTCCAAGGAAGTTGTACCGTCTTTG GCCTGTGAATGGTCTCAACTAACCCTTTCAGGTCTAAATGGAGCCCAGATGGAGAA AATACCCCTATTGCATATTTCTTCATGTGACCAAAATATTTCAGAAAAAGACCTATTA GACACAGAGAACAAAGAAGAAGAATTTTCTTACTTCAGAGAATTCTTTGCCACGT ATTTCTAGCCTACCAAAATCAGAGAAGCCATTAAATGAGGAAACAGTGGTAAATAA GAGAGATGAAGACACATCTTGAATCTCATACAGACTGCATTCTTGCAGTAAAGC AGGCAATATCTGGAACTTCTCCAGTGGCTTCTTCATTTCAGGGTATCAAAAAGTCTA TATTCAGAATAAGAGAATCACCTAAAGAGACTTTCAATGCAAGTTTTTCAGGTCATA TGACTGATCCAAACTTTAAAAAAGAAACTGAAGCCTCTGAAAGTGGACTGGAAATA CATACTGTTTGCTCACAGAAGGAGGACTCCTTATGTCCAAATTTAATTGATAATGGA AGCTGGCCAGCCACCACACAGAATTCTGTAGCTTTGAAGAATGCAGGTTTAAT ATCCACTTTGAAAAAGAAAACAAATAAGTTTATTTATGCTATACATGATGAAACATCT TATAAAGGAAAAAAATACCGAAAGACCAAAAATCAGAACTAATTAACTGTTCAGCC CAGTTTGAAGCAAATGCTTTTGAAGCACCACTTACATTTGCAAATGCTGATTCAGGt acctctgtct

## Exon 11

tttgtgtttttatgtttagGTTTATTGCATTCTTCTGTGAAAAGAAGCTGTTCACAGAATGATTCT GĂĂGAAČCAĂCTTTGTCCTTAACTAGCTCTTTTGGGACAATTCTGAGGAAATGTTCT AGAAATGAAACATGTTCTAATAATACAGTAATCTCTCAGGATCTTGATTATAAAGAA GCAAAATGTAATAAGGAAAAACTACAGTTATTTATTACCCCAGAAGCTGATTCTCTG TCATGCCTGCAGGAAGGACAGTGTGAAAATGATCCAAAAAGCAAAAAGTTTCAGA TATAAAAGAAGAGGTCTTGGCTGCAGCATGTCACCCAGTACAACATTCAAAAGTGG AATACAGTGATACTGACTTTCAATCCCAGAAAAGTCTTTTATATGATCATGAAAATG CCAGCACTCTTATTTTAACTCCTACTTCCAAGGATGTTCTGTCAAACCTAGTCATGA TTTCTAGAGGCAAAGAATCATACAAAATGTCAGACAAGCTCAAAGGTAACAATTATG CTTTAAATGAAAATTATAAAAACGTTGAGCTGTTGCCACCTGAAAAATACATGAGAG TAGCATCACCTTCAAGAAAGGTACAATTCAACCAAAACACAAATCTAAGAGTAATCC AAAAAAATCAAGAAGAAACTACTTCAATTTCAAAAATAACTGTCAATCCAGACTCTG AAGAACTTTCTCAGACAATGAGAATAATTTTGTCTTCCAAGTAGCTAATGAAAGGA ATAATCTTGCTTTAGGAAATACTAAGGAACTTCATGAAACAGACTTGACTTGTGTAA ACGAACCCATTTTCAAGAACTCTACCATGGTTTTATATGGAGACACAGGTGATAAAC AAGCAACCCAAGTGTCAATTAAAAAAGATTTGGTTTATGTTCTTGCAGAGGAGAAC AAAAATAGTGTAAAGCAGCATATAAAAATGACTCTAGGTCAAGATTTAAAATCGGAC ATCTCCTTGAATATAGATAAAATACCAGAAAAAAAATAATGATTACATGAACAAATGG GCAGGACTCTTAGGTCCAATTTCAAATCACAGTTTTGGAGGTAGCTTCAGAACAGC TTCAAATAAGGAAATCAAGCTCTCTGAACATAACATTAAGAAGAGCAAAATGTTCTT CAAAGATATTGAAGAACAATATCCTACTAGTTTAGCTTGTGTGAAATTGTAAATAC CTTGGCATTAGATAATCAAAAGAAACTGAGCAAGCCTCAGTCAATTAATACTGTATC TGCACATTTACAGAGTAGTGTAGTTGTTCTGATTGTAAAAATAGTCATATAACCCC TCAGATGTTATTTCCAAGCAGGATTTTAATTCAAACCATAATTTAACACCTAGCCAA AAGGCAGAAATTACAGAACTTTCTACTATATTAGAAGAATCAGGAAGTCAGTTTGAA TTTACTCAGTTTAGAAAACCAAGCTACATATTGCAGAAGAGTACATTTGAAGTGCCT GAAAACCAGATGACTATCTTAAAGACCACTTCTGAGGAATGCAGAGATGCTGATCT 

# Figure 2C

AAGGTACAGTTGAAATTAAACGGAAGTTTGCTGGCCTGTTGAAAAATGACTGTAAC AAAAGTGCTTCTGGTTATTTAACAGATGAAAATGAAGTGGGGTTTAGGGGCTTTTAT TCTGCTCATGGCACAAAACTGAATGTTTCTACTGAAGCTCTGCAAAAAGCTGTGAA ACTGTTTAGTGATATTGAGAATATTAGTGAGGAAACTTCTGCAGAGGTACATCCAAT AAGTTTATCTTCAAGTAAATGTCATGATTCTGTTGTTTCAATGTTTAAGATAGAAAAT CATAATGATAAAACTGTAAGTGAAAAAAAAAATAATAAATGCCAACTGATATTACAAAATA ATATTGAAATGACTACTGGCACTTTTGTTGAAGAAATTACTGAAAATTACAAGAGAA ATACTGAAAATGAAGATAACAAATATACTGCTGCCAGTAGAAATTCTCATAACTTAG AATTTGATGGCAGTGATTCAAGTAAAAATGATACTGTTTGTATTCATAAAGATGAAA CGGACTTGCTATTTACTGATCAGCACAACATATGTCTTAAATTATCTGGCCAGTTTA TGAAGGAGGAAACACTCAGATTAAAGAAGATTTGTCAGATTTAACTTTTTTGGAAG TTGCGAAAGCTCAAGAAGCATGTCATGGTAATACTTCAAATAAAGAACAGTTAACT GCTACTAAAACGGAGCAAAATATAAAAGATTTTGAGACTTCTGATACATTTTTTCAG ACTGCAAGTGGGAAAAATATTAGTGTCGCCAAAGAGTCATTTAATAAAATTGTAAAT TTCTTTGATCAGAAACCAGAAGAATTGCATAACTTTTCCTTAAATTCTGAATTACATT CTGACATAAGAAAGAACAAAATGGACATTCTAAGTTATGAGGAAACAGACATAGTT AAACACAAAATACTGAAAGAAGTGTCCCAGTTGGTACTGGAAATCAACTAGTGAC CTTCCAGGGACAACCCGAACGTGATGAAAAGATCAAAGAACCTACTCTGTTGGGTT TTCATACAGCTAGCGGGAAAAAAGTTAAAATTGCAAAGGAATCTTTGGACAAAGTG AAAAACCTTTTTGATGAAAAAGAGCAAGGTACTAGTGAAATCACCAGTTTTAGCCAT CAATGGGCAAAGACCCTAAAGTACAGAGGGCCTGTAAAGACCTTGAATTAGCAT GTGAGACCATTGAGATCACAGCTGCCCCAAAGTGTAAAGAAATGCAGAATTCTCTC AATAATGATAAAAACCTTGTTTCTATTGAGACTGTGGTGCCACCTAAGCTCTTAAGT GATAATTTATGTAGACAAACTGAAAATCTCAAAAACATCAAAAAGTATCTTTTTGAAAG TTAAAGTACATGAAAATGTAGAAAAAGAAAAGCAAAAAAGTCCTGCAACTTGTTACA CAAATCAGTCCCTTATTCAGTCATTGAAAATTCAGCCTTAGCTTTTTACACAAGTT GTAGTAGAAAAACTTCTGTGAGTCAGACTTCATTACTTGAAGCAAAAAAATGGCTTA GAGAAGGAATATTTGATGGTCAACCAGAAAGAATAAATACTGCAGATTATGTAGGA AATTATTTGTATGAAAATAATTCAAACAGTACTATAGCTGAAAAATGACAAAAATCATC TCTCCGAAAACAAGATACTTATTTAAGTAACAGTAGCATGTCTAACAGCTATTCCT ACCATTCTGATGAGGTATATAATGATTCAGGATATCTCTCAAAAAATAAACTTGATT CTGGTATTGAGCCAGTATTGAAGAATGTTGAAGATCAAAAAAACACTAGTTTTTCCA AAGTAATATCCAATGTAAAAGATGCAAATGCATACCCACAAACTGTAAATGAAGATA TTTGCGTTGAGGAACTTGTGACTAGCTCTTCACCCTGCAAAAATAAAAATGCAGCC ATTAAATTGTCCATATCTAATAGTAATAATTTTGAGGTAGGGCCACCTGCATTTAGG ATAGCCAGTGGTAAAATCGTTTGTGTTTCACATGAAACAATTAAAAAAGTGAAAGAC ATATTTACAGACAGTTTCAGTAAAGTAATTAAGGAAAACAACGAGAATAAATCAAAA ATTTGCCAAACGAAAATTATGGCAGGTTGTTACGAGGCATTGGATGATTCAGAGGA TATTCTTCATAACTCTCTAGATAATGATGAATGTAGCACGCATTCACATAAGGTTTTT GCTGACATTCAGAGTGAAGAAATTTTACAACATAACCAAAATATGTCTGGATTGGA GAAAGTTTCTAAAATATCACCTTGTGATGTTAGTTTGGAAACTTCAGATATATGTAAA TGTAGTATAGGGAAGCTTCATAAGTCAGTCTCATCTGCAAATACTTGTGGGATTTTT AGCACAGCAAGTGGAAAATCTGTCCAGGTATCAGATGCTTCATTACAAAACGCAAG ACAAGTGTTTCTGAAATAGAAGATAGTACCAAGCAAGTCTTTTCCAAAGTATTGTT CTCCAGAACATTTAATATCCCAAAAAGGCTTTTCATATAATGTGGTAAATTCATCTG

# Figure 2D

CTTTCTCTGGATTTAGTACAGCAAGTGGAAAGCAAGTTTCCATTTTAGAAAGTTCCT TACACAAAGTTAAGGGAGTGTTAGAGGAATTTGATTTAATCAGAACTGAGCATAGT CTTCACTATTCACCTACGTCTAGACAAAATGTATCAAAAATACTTCCTCGTGTTGAT AAGAGAAACCCAGAGCACTGTGTAAACTCAGAAATGGAAAAAACCTGCAGTAAAGA ATTTAAATTATCAAATAACTTAAATGTTGAAGGTGGTTCTTCAGAAAATAATCACTCT ATTAAAGTTTCTCCATATCTCTCAATTTCAACAAGACAACAACAGTTGGTATTAG GAACCAAAGTCTCACTTGTTGAGAACATTCATGTTTTGGGAAAAGAACAGGCTTCA CCTAAAAACGTAAAAATGGAAATTGGTAAAACTGAAACTTTTTCTGATGTTCCTGTG AAAACAAATATAGAAGTTTGTTCTACTTACTCCAAAGATTCAGAAAACTACTTTGAAA CAGAAGCAGTAGAAATTGCTAAAGCTTTTATGGAAGATGATGAACTGACAGATTCT AAACTGCCAAGTCATGCCACACATTCTCTTTTTACATGTCCCGAAAATGAGGAAATG GTTTTGTCAAATTCAAGAATTGGAAAAAGAAGAGGGGGGGCCCCTTATCTTAGTGGgt aagtgttcatttttacctttcgtgttgccaatca

# Exon 12

aaaacatatatgaaatatttctttttagGAGAACCCTCAATCAAAAGAAACTTATTAAATGAATTTG ACAGGATAATAGAAAATCAAGAAAAATCCTTAAAGGCTTCAAAAAGCACTCCAGAT Ggtaaaattagctttttattata

# Exon 13

aatatgtaatataaaataattgtttcctagGCACAATAAAAGATCGAAGATTGTTTATGCATCATGT TTCTTTAGAGCCGATTACCTGTGTACCCTTTCGgtaagacatgtttaaatttttctaa

## Exon 14

ccccattgcagCACAACTAAGGAACGTCAAGAGATACAGAATCCAAATTTTACCGCACC TGGTČAAĞAATTTCTGTCTAAATCTCATTTGTATGAACATCTGACTTTGGAAAAATCT TCAAGCAATTTAGCAGTTTCAGGACATCCATTTTATCAAGTTTCTGCTACAAGAAAT GAAAAATGAGACACTTGATTACTACAGGCAGACCAAACCAAAGTCTTTGTTCCACC TTTTAAAACTAAATCACATTTTCACAGAGTTGAACAGTGTGTTAGGAATATTAACTTG GAGGAAAACAGACAAAGCAAAACATTGATGGACATGGCTCTGATGATAAAAAA TAAGATTAATGACAATGAGATTCATCAGTTTAACAAAAACAACTCCAATCAAGCAGC  ${\tt AGCTGTAACTTTCACAAAGTGTGAAGAAGAACCTTTAGgtattgtatgacaatttgtgtgatgaatt}$ 

#### Exon 15

tttttgctaagtatttattctttgatagATTTAATTACAAGTCTTCAGAATGCCAGAGATATACAGGAT ATĞCGĂATTAAGĂAGĂAACAAAGGCAACGCGTCTTTCCACAGCCAGGCAGTCTGTA TCTTGCAAAAACATCCACTCTGCCTCGAATCTCTCTGAAAGCAGCAGTAGGAGGCC AAGTTCCCTCTGCGTGTTCTCATAAACAGgtatgtgt

## Exon 16

tttttcttttttgtgtgtgtttattttgtgtagCTGTATACGTATGGCGTTTCTAAACATTGCATAAAAATTAACAGCĂĂĂĂTGCĂĞAĞTCTTTTCAGTTTCACACTGAAGATTATTTTGGTAAGGAAA GTTTATGGACTGGAAAAGGAATACAGTTGGCTGATGGTGGATGGCTCATACCCTCC AATGATGGAAAGGCTGGAAAAGAAGAATTTTATAGgtactctatgcaaaaagattgtgtgttaactttt atg

# Figure 2E

Exon 17

ttatttgttcagGGCTCTGTGTGACACTCCAGGTGTGGATCCAAAGCTTATTTCTAGAATTT GGĞTTTĂTAATCACTATAGATGGATCATATGGAAACTGGCAGCTATGGAATGTGCC TTTCCTAAGGAATTTGCTAATAGATGCCTAAGCCCAGAAAGGGTGCTTCTTCAACTA AAATACAGgcaagtttaaagcatt

Exon 18

ttttgttttcacttttagATATGATACGGAAATTGATAGAAGCAGAAGATCGGCTATAAAAAAAGA TAATGGAAAGGGATGACACAGCTGCAAAAACACTTGTTCTCTGTGTTTCTGACATA ATTTCATTGAGCGCAAATATATCTGAAACTTCTAGCAATAAAACTAGTAGTGCAGAT ACCCAAAAGTGGCCATTATTGAACTTACAGATGGGTGGTATGCTGTTAAGGCCCA GTTAGATCCTCCCCTCTTAGCTGTCTTAAAGAATGGCAGACTGACAGTTGGTCAGA AGATTATTCTTCATGGAGCAGAACTGGTGGGCTCTCCTGATGCCTGTACACCTCTT GAAGCCCCAGAATCTCTTATGTTAAAGgtaaatt

Exon 19

taaatcaatatttattaatttgtccagATTTCTGCTAACAGTACTCGGCCTGCTCGCTGGTATAC CAAACTTGGATTCTTTCCTGACCCTAGACCTTTTCCTCTGCCCTTATCATCGCTTTT CAGTGATGGAGGAAATGTTGGTTGTTGATGTAATTATTCAAAGAGCATACCCTAT ACAGgtatgatgtattcttgaaactta

Exon 20

tttggtgtgtgtaacacattattacagTGGATGGAGAAGACATCATCTGGATTATACATATTTCGCAĂŤĞĂĂĞAĞAĞĞAAĞĂAAAĞĞAAĞCAĞCAAAATATĞTĞĞAĞĞCCCAACAAAĞĀ GACTAGAAGCCTTATTCACTAAAATTCAGGAGGAATTTGAAGAACATGAAGgtaaaatt agttatatggtacacattgttatttc

Exon 21

agtttagtgaattaataatccttttgttttcttagAAAACACAACAAAACCATATTTACCATCACGTGCAC TĂACĂĂGACAGCAAGŤTCGTĞCTTTGCAAGATGGTGCAGAGCTTTATGAAGCAGTG AAGAATGCAGCAGACCCAGCTTACCTTGAGgtgagaggtaagaggacatataatgag

Exon 22

tttttattccaatatcttaaatggtcacagGGTTATTTCAGTGAAGAGCAGTTAAGAGCCTTGAATAA TCACAGGCAAATGTTGAATGATAAGAAACAAGCTCAGATCCAGTTGGAAATTAGGA AGgCCATGGAATCTGCTGAACAAAAGGAACAAGGTTTATCAAGGGATGTCACAACC GTGTGGAAGTTGCGTATTGTAAGCTATTCAAAAAAAAGAAAAAGATTCAGgtaagtatgta aatgctttgttttta

Exon 23

tctccaaacagTTATACTGAGTATTTGGCGTCCATCATCAGATTTATATTCTCTGTTAACA GAAGGAAAGGAATTTATCATCTTGCAACTTCAAAATCTAAAAGTAAATCT GAAAGAGCTAACATACAGTTAGCAGCGACAAAAAAAACTCAGTATCAACAACTACC Gotacaaacctttcattgtaattttt

# Figure 2F

Exon 24

gaatttttgttttctgtagGTTTCAGATGAAATTTTATTTCAGATTTACCAGCCACGGGAGC CCCTTCACTTCAGCAAATTTTTAGATCCAGACTTTCAGCCATCTTGTTCTGAGGTGG ACCTAATAGGATTTGTCGTTTCTGTTGAAAAAAACAGgtaatgcacaatatagttaatttttttat tgattcttttaaaaaaacattgtct

Exon 25

taacattcttttcttttttttccattctagGACTTGCCCCTTTCGTCTATTTGTCAGACGAATGTTACAA TTTACTGGCAATAAAGTTTTGGATAGACCTTAATGAGGACATTATTAAGCCTCATAT GTTAATTGCTGCAAGCAACCTCCAGTGGCGACCAGAATCCAAATCAGGCCTTCTTA CTTTATTTGCTGGAGATTTTTCTGTGTTTTCTGCTAGTCCAAAAGAGGGCCACTTTC AAGAGACATTCAACAAAATGAAAAATACTGTTGAGgtaaggtta

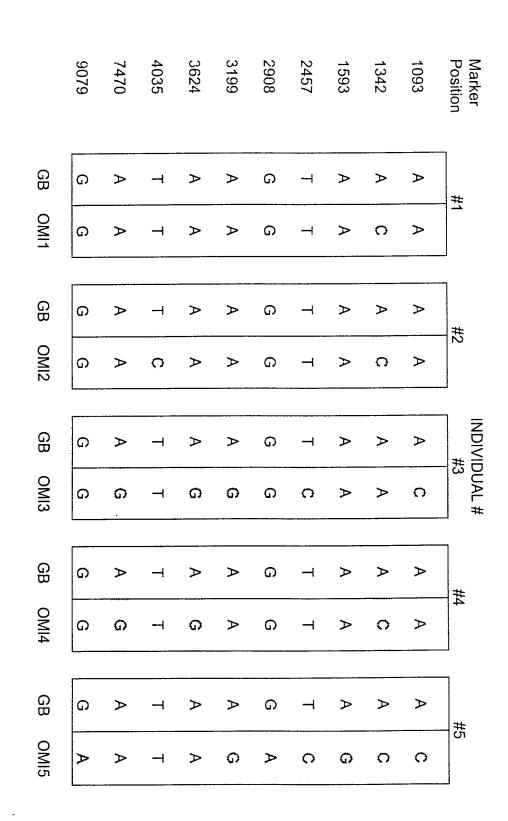
Exon 26

ataaagcagcttttccacttattttcttagAATATTGACATACTTTGCAATGAAGCAGAAAACAAGCT TATGCATATACTGCATGCĂAATGATCCCAAGTGGTCCACCCCAACTAAAGACTGTA CTTCAGGGCCGTACACTGCTCAAATCATTCCTGGTACAGGAAACAAGCTTCTGgtaa gttaatgtaaactcaaggaatattataag

Exon 27

tacgttttcatttttttatcagATGTCTTCTCCTAATTGTGAGATATATTATCAAAGTCCTTTATCA CTTTGTATGGCCAAAAGGAAGTCTGTTTCCACACCTGTCTCAGCCCAGATGACTTC AAAGTCTTGTAAAGGGGAGAAAGAGATTGATGACCAAAAGAACTGCAAAAAGAGAA GAGCCTTGGATTTCTTGAGTAGACTGCCTTTACCTCCACCTGTTAGTCCCATTTGTA CATTTGTTTCTCCGGCTGCACAGAAGGCATTTCAGCCACCAAGGAGTTGTGGCAC CAAATACGAAACACCCATAAAGAAAAAAGAACTGAATTCTCCTCAGATGACTCCATT TAAAAAATTCAATGAAATTTCTCTTTTGGAAAGTAATTCAATAGCTGACGAAGAACTT GCATTGATAAATACCCAAGCTCTTTTGTCTGGTTCAACAGGAGAAAAAACAATTTATA TCTGTCAGTGAATCCACTAGGACTGCTCCCACCAGTTCAGAAGATTATCTCAGACT GAAACGACGTTGTACTACATCTCTGATCAAAGAACAGGAGAGTTCCCAGGCCAGTA CGGAAGAATGTGAGAAAAATAAGCAGGACACAATTACAACTAAAAAATATATCTAA GCATTTGCAAAGGCGACAATAAATTATTGACGCTTAACCTTTCCAGTTTATAAGACT **GGA** 

# FIGURE 3



Docket No:	
05371.31.US02	
PA-	
	05371.31.US02

Applicant or Patentee:

Patricia D. Murphy; Marga B. White; Mark B. Rabin; Sheri J. Olson; Matthew Yoshikawa; Geoffrey

M. Jackson; Tara Eskandari; Brenda Schryer; and Michael Park.

Serial or Patent No.:

To be assigned.

Filed or Issued:

Herewith

Title:

NOVEL CODING SEQUENCE HAPLOTYPES OF THE HUMAN BRCA2 GENE

I hereby declare that I am

the owner of the small business concern identified below:

XXX an official of the small business concern empowered to act on behalf of the concern identified below:

NAME OF SMALL BUSINESS CONCERN:

Oncormed, Inc. 205 Perry Parkway

ADDRESS OF SMALL BUSINESS CONCERN:

Gaithersburg, MD 20877

I hereby declare that the above-identified small business concern qualifies as a small business concern as defined in 13 CFR 121.12, and reproduced in 37 CFR 1.9(d), for purposes of paying reduced fees to the United States Patent and Trademark Office, in that the number of employees of the concern, including those of its affiliates, does not exceed 500 persons. For purposes of this statement, (1) the number of employees of the business concern is the average over the pervious fiscal year of the concern of the persons employed on a full-time, part-time or temporary basis during each of the pay periods of the fiscal year, and (2) concerns are affiliates of each other when either, directly or indirectly, one concern controls of has the power to control the other, or a third party or parties controls or has the power to control both.

I hereby declare that rights under contract or law have been conveyed to and remain with the small business concern identified above with regard to the invention described in:

XXXX_	the specification filed herewith with title as listed above.
	the application identified above.
	the patent identified above.

If the rights held by the above identified small business concern are not exclusive, each individual, concern or organization having rights in the invention must file separate verified statements averring to their status as small entities, and no rights to the invention are held by any person, other than the inventor, who would not qualify as an independent inventor under 37 CFR 1.9(c) if that person made the invention, or by any concern which would not qualify as a small business concern under 37 CFR 1.9(d), or a nonprofit organization under 37 CFR 1.9(e).

	Each person, concern or organization having any rights in the invention is listed below
XXX	No such person, concern or organization exists.
	Each such person, concern or organization is listed below.

Separate verified statements are required from each named person, concern or organization having rights to the invention averring to their status as small entities. (37 CFR 1.27)

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fees due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b)).

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, and patent issuing thereon, or any patent to which this verified statement is directed.

NAME OF PERSON SIGNING:

TITLE OF PERSON IF OTHER THAN OWNER: ADDRESS OF PERSON SIGNING:

DOUG DOLGINOW, M.D. President & C.O.O.

205 Perry Parkway Gaithersburg, MD 20877

DATE 5/13/98

# Combined Declaration and Power of Attorney for Patent Application

Docket Number: 5371.31.US02

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am an original, first and joint inventor (if plural names are listed below) of the subject matter that is claimed and for which a patent is sought on the invention entitled **NOVEL CODING SEQUENCE HAPLOTYPES OF THE HUMAN BRCA2 GENE**, the specification of which is attached hereto unless the following box is checked:

was filed on	Herewith		;			
as United States	Application	Number or PCT I	International A	pplication Number	 To Be Assigned	_; and
was amended on	1		(if applicable).			

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information that is material to patentability as defined in 37 C.F.R. § 1.56.

I hereby claim foreign priority benefits under 35 U.S.C. § 119(a)-(d) or § 365(b) of any foreign application(s) for patent or inventor's certificate, or § 365(a) of any PCT international application, which designated at least one country other than the United States listed below, and have also identified below any foreign application for patent or inventor's certificate, or PCT international application having a filing date before that of the application on which priority is claimed.

Application No.	Country	(Day/Month/Year/Filed)	Priority Claimed
	•		Yes No
			Yes No
			Yes No
			Yes No

I hereby claim the benefit under 35 U.S.C. §119(e) of any United States provisional application(s) listed below.

Application No.	Filing Date
60/055,784	August 15, 1997
60/064,926	November 7, 1997
60/065,367	November 12, 1997

I hereby claim the benefit under 35 U.S.C. § 120 of any United States application(s), or under § 365(c) of any PCT international application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT international application in the manner provided by the first paragraph of 35 U.S.C. § 112, I acknowledge the duty to disclose information that is material to patentability as defined in 37 C.F.R. § 1.56 that became available between the filing date of the prior application and the national or PCT international filing date of this application.

Application No.	Filing Date	(Status – patented, pending, abandoned)

I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith:

Jeffrey I. Auerbach, Reg. No. 32,680 Melvin L. Barnes, Jr. Reg. No. 38,375 Thane Bauz, Reg. No. P41,604 Michael J. Bell, Reg. No. 39,604 John A. Bendrick, Reg. No. P41,612 Mark R. Buscher, Reg. No. 35,006 Celine T. Callahan, Reg. No. 34,301 Cono A. Carrano, Reg. No. 39,623 James F. Davis, Reg. No. 21,072 Thomas M. Dunham, Reg. No. 39,965 Joel M. Freed, Reg. No. 25,101 Vernon Randall Gard, Reg. No. 33,886

Alan M. Grimaldi, Reg. No. 26,599 Alexander J. Hadjis, Reg. No. 36,540 Albert P. Halluin, Reg. No. 25,227 Michael N. Haynes, Reg. No. 40,014 Rouget F. Henschel, Reg. No. 39,221 Leslie L. Jacobs, Jr., Reg. No. 40,659 Richard H. Kjeldgaard, Reg. No. 30,186 Joseph P. Lavelle, Reg. No. 31,036 David R. Marsh, Reg. No. 41,408 Kevin W. McCabe, Reg. No. 41,182 Joseph A. Micallef, Reg. No. 39,772 Anthony D. Miller, Reg. No. 34,394 Karen L. Nicastro, Reg. No. 35,968
Bradley J. Olson, Reg. No. 40,750
Russell O. Paige, Reg. No. P40,758
Stephen J. Pentlicki, Reg. No. 40,125
Andrew Y. Piatnicia, Reg. No. 40,772
Andrea G. Reister, Reg. No. 36,253
Stephen J. Rosenman, Reg. No. 29,209
David P. Ruschke, Reg. No. 40,151
Timothy L. Scott, Reg. No. 37,931
Anthony W. Shaw, Reg. No. 37,931
Anthony W. Shaw, Reg. No. 39,839
Michael J. Songer, Reg. No. 39,841
R. Thomas Gallages
Reg. No. 32, 692
John E. Tarcga
Reg. No. 33, 683

Send Correspondence to:

## Albert P. Halluin HOWREY & SIMON

Box No. 34 1299 Pennsylvania Avenue, N.W. Washington, D.C. 20004-2402 Facsimile: (202) 383-7195

Direct Telephone Calls to: (202) 783-0800

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

FULL NAME OF SOLE OR FIRST INVENTOR:	CITIZENSHIP:
Patricia D. Murphy	USA
RESIDENCE:	DATE:
Slingerlands, New York, USA	578/98
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	Samuel Murphy
	J
FULL NAME OF SECOND INVENTOR:	CITIZENSHIP:
Marga B. White	USA
RESIDENCE:	DATE:
Frederick, Maryland, USA	5,13/95
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	Mary Outh Come
FULL NAME OF THIRD INVENTOR:	CITIZENSHIP:
Mark B. Rabin	USA
RESIDENCE:	DATE:_/
Rockville, Maryland, USA	5/13/98
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	Make tile

(Supply similar information and signature for subsequent joint inventors, if any)

FULL NAME OF FOURTH INVENTOR:	CITIZENSHIP
Sheri J. Olson	USA ( (L)
RESIDENCE:	DATE: 5/13/98
Falls Church, Virginia, USA	
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	1 1 WW ADN LA CON.
FULL NAME OF FIFTH INVENTOR:	CITIZENSHIP:
Matthew Yoshikawa	USA
RESIDENCE ADDRESS:	DATE:
Germantown, Maryland, USA	
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	
FULL NAME OF SIXTH INVENTOR:	CITIZENSHIP:
Geoffrey M. Jackson	USA
RESIDENCE ADDRESS:	DATE:
Beltsville, Maryland, USA	
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	
FULL NAME OF SEVENTH INVENTOR:	CITIZENSHIP:
Tara Eskandari	USA
RESIDENCE ADDRESS:	DATE: 5-13-98
Rockville, Maryland, USA	3-13-10
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	INVENTOR'S SIGNATURE:
FULL NAME OF EIGHTH INVENTOR:	CITIZENSHIP:
Brenda Schryer	USA
RESIDENCE ADDRESS:	DATE: 13 OF
Bel Air, Maryland, USA	5-16-98
POST OFFICE ADDRESS:	
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	inventor's signature: Rooidah Schuler
	J
FULL NAME OF NINTH INVENTOR:	CITIZENSHIP:
Michael Park	USA
RESIDENCE ADDRESS:	DATE: 5/13/98
Rockville, Maryland, USA	
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE A CONT
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	Concount of the

# Combined Declaration and Power of Attorney for Patent Application

Docket Number: 5371.31.US02

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am an original, first and joint inventor (if plural names are listed below) of the subject matter that is claimed and for which a patent is sought on the invention entitled **NOVEL CODING SEQUENCE HAPLOTYPES OF THE HUMAN BRCA2 GENE**, the specification of which is attached hereto unless the following box is checked:

was filed on	Herewith	;	
as United States	s Application	Number or PCT International Application Number	To Be Assigned; and
was amended o	n	(if applicable).	

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information that is material to patentability as defined in 37 C.F.R. § 1.56.

I hereby claim foreign priority benefits under 35 U.S.C. § 119(a)-(d) or § 365(b) of any foreign application(s) for patent or inventor's certificate, or § 365(a) of any PCT international application, which designated at least one country other than the United States listed below, and have also identified below any foreign application for patent or inventor's certificate, or PCT international application having a filing date before that of the application on which priority is claimed.

Application No.	Country	(Day/Month/Year/Filed)	Priority Claimed
			Yes No

I hereby claim the benefit under 35 U.S.C. §119(e) of any United States provisional application(s) listed below.

Application No.	Filing Date
60/055,784	August 15, 1997
60/064,926	November 7, 1997
60/065,367	November 12, 1997

I hereby claim the benefit under 35 U.S.C. § 120 of any United States application(s), or under § 365(c) of any PCT international application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT international application in the manner provided by the first paragraph of 35 U.S.C. § 112, I acknowledge the duty to disclose information that is material to patentability as defined in 37 C.F.R. § 1.56 that became available between the filing date of the prior application and the national or PCT international filing date of this application.

Application No.	Filing Date	(Status – patented, pending, abandoned)

I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith:

Jeffrey I. Auerbach, Reg. No. 32,680 Melvin L. Barnes, Jr. Reg. No. 38,375 Thane Bauz, Reg. No. P41,604 Michael J. Bell, Reg. No. 39,604 John A. Bendrick, Reg. No. P41,612 Mark R. Buscher, Reg. No. 35,006 Celine T. Callahan, Reg. No. 34,301 Cono A. Carrano, Reg. No. 39,623 James F. Davis, Reg. No. 21,072 Thomas M. Dunham, Reg. No. 39,965 Joel M. Freed, Reg. No. 25,101 Vernon Randall Gard, Reg. No. 33,886

Send Correspondence to:

Alan M. Grimaldi, Reg. No. 26,599
Alexander J. Hadjis, Reg. No. 36,540
Albert P. Halluin, Reg. No. 25,227
Michael N. Haynes, Reg. No. 40,014
Rouget F. Henschel, Reg. No. 39,221
Leslie L. Jacobs, Jr., Reg. No. 40,659
Richard H. Kjeldgaard, Reg. No. 30,186
Joseph P. Lavelle, Reg. No. 31,036
David R. Marsh, Reg. No. 41,408
Kevin W. McCabe, Reg. No. 41,182
Joseph A. Micallef, Reg. No. 39,772
Anthony D. Miller, Reg. No. 34,394

# Albert P. Halluin HOWREY & SIMON

Box No. 34 1299 Pennsylvania Avenue, N.W. Washington, D.C. 20004-2402 Facsimile: (202) 383-7195 Karen L. Nicastro, Reg. No. 35,968
Bradley J. Olson, Reg. No. 40,750
Russell O. Paige, Reg. No. P40,758
Stephen J. Pentlicki, Reg. No. 40,125
Andrew Y. Piatnicia, Reg. No. 40,772
Andrea G. Reister, Reg. No. 36,253
Stephen J. Rosenman, Reg. No. 29,209
David P. Ruschke, Reg. No. 40,151
Timothy L. Scott, Reg. No. 37,931
Anthony W. Shaw, Reg. No. 30,104
J. David Smith, Reg. No. 39,839
Michael J. Songer, Reg. No. 39,841
R. Thomas Galleges
Reg. No. 32, 692
John E. Tarcga
Reg. No. 33,683

Direct Telephone Calls to: (202) 783-0800

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

FULL NAME OF SOLE OR FIRST INVENTOR:	CITIZENSHIP:
Patricia D. Murphy	USA
RESIDENCE:	DATE:
Slingerlands, New York, USA	578/98
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	Satrua & murphy
FULL NAME OF SECOND INVENTOR:	CITIZENSHIP:
Marga B. White	USA
RESIDENCE:	DATE:
Frederick, Maryland, USA	5/13/98
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	marge but which
FULL NAME OF THIRD INVENTOR:	CITIZENSHIP:
Mark B. Rabin	USA
RESIDENCE:	DATE:/
Rockville, Maryland, USA	5/13/98
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	Illak Kile

(Supply similar information and signature for subsequent joint inventors, if any)

FULL NAME OF FOURTH INVENTOR:	CITIZENSHIP
Sheri J. Olson	USA
RESIDENCE:	DATE: 5/12/90
Falls Church, Virginia, USA	5/13/98
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	VIDA AON CISON.
FULL NAME OF FIFTH INVENTOR:	CITIZENSHIP:
Matthew Yoshikawa	USA
RESIDENCE ADDRESS:	DATE: 1 1
Germantown, Maryland, USA	1 3 UB 17 Z
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	Redden Verstrle
FULL NAME OF SIXTH INVENTOR:	CITIZENSHIP:
Geoffrey M. Jackson	USA
RESIDENCE ADDRESS:	DATE:
Beltsville, Maryland, USA	5/15/78
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	I to forh Son
, , , , , , , , , , , , , , , , , , , ,	
FULL NAME OF SEVENTH INVENTOR:	CITIZENSHIP:
Tara Eskandari	USA
RESIDENCE ADDRESS:	DATE: 5-13-98
Rockville, Maryland, USA	3-13-10
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	INVENTOR'S SIGNATURE:
FULL NAME OF EIGHTH INVENTOR:	CITIZENSHIP:
Brenda Schryer	USA
RESIDENCE ADDRESS:	DATE:
Bel Air, Maryland, USA	
POST OFFICE ADDRESS:	INVENTOR'S SIGNATURE:
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	
, , , , , , , , , , , , , , , , , , , ,	
FULL NAME OF NINTH INVENTOR:	CITIZENSHIP:
Michael Park	USA
RESIDENCE ADDRESS:	DATE: 5/13/98
Rockville, Maryland, USA	5/15/10
POST OFFICE ADDRESS:	INVENTORIES SIGNATURE OF CONT
205 Perry Parkway, Gaithersburg, Maryland 20877, USA	Concorne p. 1 an